

## GRAVITATIONAL INSTABILITY IN TWO-PHASE DISKS AND THE ORIGIN OF THE MOON

CHRISTOPHER THOMPSON

Joseph Henry Laboratories, Princeton University

AND

DAVID J. STEVENSON

Division of Geological and Planetary Sciences, California Institute of Technology

Received 1986 August 28; accepted 1988 March 22

### ABSTRACT

Two-phase disks may be gravitationally unstable at temperatures or surface densities at which a disk composed of either single phase would be highly stable. We argue that two-phase disks can achieve a marginally unstable state (in addition to a highly unstable state that leads to fragmentation), limited by the ability of the photosphere to radiate the energy dissipated in the disk. We provide a self-consistent prescription for the viscosity induced by the slow instabilities. Two-phase disks are more centrally condensed than single-phase disks, and their secular cooling time may be comparable to their spreading time. A circumterrestrial disk of sufficient mass to form the Moon provides a detailed example of all the preceding points. We investigate its stability, structure, and dynamical evolution, and conclude that its spreading time is short ( $\sim 100$  yr); the Moon is formed molten, or partially molten; the Moon's initial orbit lies in the Earth's equatorial plane; and only a small fraction of the disk mass is lost in a wind, although this may represent a substantial fraction of volatiles. Most of these conclusions are independent of how the disk was formed, e.g., from a giant impact.

*Subject headings:* Moon — planets: formation

### I. INTRODUCTION

Disks of particulate or gaseous material exist in a variety of extraterrestrial environments. Saturn's rings and accretion disks around compact stars are familiar examples. Disks are also believed to be precursors to planetary systems (i.e., the solar "nebula") and satellite systems, perhaps including the Earth's moon. In the currently favored view of solar system origin (see recent reviews by Wetherill 1980; Hayashi, Nakazawa, and Nakagawa 1985), the accumulation of planets and their satellites begins in gravitationally unstable particulate or gaseous disks.

The widest gap in our understanding of disks lies in the nature of the viscosity, which controls how fast a disk spreads and liberates its gravitational binding energy. Without listing the whole plethora of mechanisms that have been proposed to account for the viscosity, we note that they often depend on the existence of some instability in the disk. In order to obtain significant spreading rates in systems of planetary or stellar dimensions, one usually requires a viscosity many orders of magnitude higher than provided by ordinary particle collisions (except for very hot plasmas where the mean free path becomes comparable to the size of the system). In this paper, we show how one type of instability, due to the self-gravity of the disk, may occur much more easily than has previously been suspected.

It has not been demonstrated that disks are stable *under any circumstances*, but a linear stability analysis predicts instabilities when the surface mass density or the compressibility of disk material is high enough. Basically, pressure stabilizes disks to short length-scale perturbations, and rotation to large length-scale perturbations. On intermediate length-scales the gravitational self-attraction of a segment of the disk may dominate these stabilizing effects, if the sound speed or velocity dispersion is low enough, or the surface mass density is high enough. Even if the disk surface density and the central potential are fixed, the stabilizing effects of pressure can vary enormously, depending on the thermal and compositional state of the disk material. Here we focus on the implications of the very low sound speed which characterizes two-phase media. A familiar example of such a medium is the frothy "head" of a beer (see Aubert, Kraynik, and Rand 1986).

In a two-phase medium, where one of the phases is vapor, the sound speed can be much lower than that in either pure phase at identical temperature and pressure. We identify two cases: the two phases can be different substances or have a large component of the same substance. In both cases, the medium may be as compressible as a gas and yet have a density approaching that of a liquid ( $\sim 1 \text{ g cm}^{-3}$ ). An example of a two-component, two-phase medium is a frothy air-water mixture: at 1 bar pressure, with an air mass fraction of  $10^{-3}$ , the sound speed is  $\sim 20 \text{ m s}^{-1}$  compared with  $1460 \text{ m s}^{-1}$  for pure water (Kieffer 1977). Very porous media can also exhibit very low sound speeds, as is the case for microporous silica "aerogel," which has the lowest sound speed ever reported in an inorganic solid,  $\sim 120 \text{ m s}^{-1}$  (Gronauer and Fricke 1986).

When there is phase equilibrium, the reduction in sound speed is correspondingly greater: material changes phase (from liquid to vapor, say) upon compression of the medium, and the latent heat absorbed reduces the increase in pressure for a given increase in density. We will concentrate on single-component media.

*If a disk has a two-phase composition, then it may be gravitationally unstable even though a single-disk would be stable for temperatures several orders of magnitude lower, or surface mass densities several orders of magnitude higher* (Thompson and Stevenson

1983). This effect can be more pronounced in the single-component case (e.g., liquid and vapor silicates) than in the two-component case (e.g., water ice particles immersed in hydrogen gas).

In particular, our work has been motivated by the importance of two-phase gravitational instabilities in understanding the formation of the Earth's moon (the Moon) from an orbiting disk. Cameron and Ward (1976) proposed that following the impact of a Mars-sized body with the proto-Earth, two or more Moon masses of silicate vapor and liquid were injected into near-Earth orbit. They originally suggested that this material must first have cooled and solidified to form a particulate disk. The disk then evolved rapidly: viscous stresses drove material beyond the Roche limit, where the Moon coalesced. In this paper, we show that a disk of liquid and vapor silicates of the requisite mass is capable of gravitational instabilities even when it is very hot. As a consequence, the disk can spread and cool on a time scale  $\sim 100$  yr. The Moon is thereby formed molten, or partially molten. This is consistent with the inferred extent of lunar differentiation (e.g., Taylor 1982).

It is not our purpose here to develop a complete model of lunar origin or to justify fully the starting state of our calculations. A broader discussion of these issues can be found in Cameron (1986), Wood (1986), and Stevenson (1987). Here we review briefly the main questions addressed and raised by the giant impact hypothesis:

1. How can we explain the large ratio of angular momentum to mass of the Earth-Moon system? If most of the Earth's mass is accreted as small particles, then it does not seem possible to supply sufficient angular momentum, except by imposing implausible constraints on the orbits of the accreting bodies (Harris 1977; but see the discussion in Wood 1986). All the angular momentum can nonetheless be supplied by a single, large oblique impact; it is this property which motivated Cameron and Ward (1976) to advocate a singular event. (In fact, the angular momentum could be supplied by a small number of large events; the arguments for a singular event are not compelling.)

2. Did large impacts occur? This question cannot be answered, except in the context of specific accretion models. However, present modeling efforts (e.g., Wetherill 1985) indicate that there should be no discontinuity between the spectrum of observed terrestrial planetary masses and the spectrum of masses from which they form. In general, the spectrum produced by a coagulation process has most of the mass in the largest bodies. Large impacts are highly probable, as Hartmann and Davis (1975) first emphasized. Indeed, there is a strong observational hint that one, or at most a few, catastrophic events determine the final magnitude and orientation of a planet's rotational angular momentum. This hint lies in the wide distribution of these angular momenta (consider Venus's slow retrograde rotation, and Uranus's high obliquity).

3. What are the consequences of a giant impact? At an expected impact velocity of  $10\text{--}14 \text{ km s}^{-1}$ , between one-half and one projectile mass of the most highly shocked material is vapor upon adiabatic release to 1 kbar (Stevenson 1985). The volume occupied by this vapor is comparable to the volume of the Earth. Pressure gradient forces may significantly deflect this heated material, which must deviate from Keplerian trajectories to reach orbit. Although a number of plausibility arguments can be advanced to support the likelihood of orbital injection (Stevenson 1984, 1985), the quantification of this process requires a three-dimensional numerical hydrodynamical calculation. A crude three-dimensional calculation, not including the shock physics of the impact, has been made by Cameron (1985) and indicates that several Moon masses may be injected into orbit. More recent calculations (Benz, Slattery, and Cameron 1986, 1987) generally support this picture, although they indicate a lower efficiency of injection.

4. Can the material which reaches orbit subsequently coalesce to form the Moon? In this paper we provide a tentative affirmative answer and outline the main implications of a disk origin for the composition and structure of the Moon.

The plan of the paper is as follows. In § II, we discuss the compressibility of two-phase media. In § III, we briefly review the stability properties of disks and show how a two-phase disk that is unstable due to its own self-gravity may enter a marginally unstable, metastable state. We outline how a disk's stability and structure are modified if it has a two-phase composition. In § IV, we discuss in detail one example of a hot, two-phase disk that should pass through the aforementioned metastable state: a circumterrestrial disk of sufficient mass to form the Moon. We reach some reasonably firm conclusions about the structure and dynamical evolution of such a disk, but the implications for the initial state and composition of the Moon are far from clear. In the Appendix, we calculate the two phase sound speed in liquid-vapor silicates.

## II. THE TWO-PHASE SOUND SPEED

Here we discuss some of the peculiar properties of a two-phase, one-component medium undergoing slight compression. If the compression occurs slowly enough, phase equilibrium is maintained and the sound speed

$$c \equiv \sqrt{\left(\frac{\partial P}{\partial \rho}\right)_s} = \frac{(\rho_2/\rho_1)(1-x) + x}{\sqrt{(C_{p1} T/l^2)(1-x) + [(\mu/kT) - (2/l) + (C_{p2} T/l^2)]x}} \quad (1)$$

Landau and Lifshitz 1959). In this expression,  $\mu$  is the molecular weight in phase 2 (which we take to be vapor),  $l$  is the heat of phase transition per unit mass, and  $C_{pi}$  is the specific heat per unit mass at constant pressure of phase  $i$ . We take phase 1 to be solid or liquid. The quantity  $x$  is the fraction of the total mass in the vapor phase.

As  $x \rightarrow 0$ ,

$$c \rightarrow \left(\frac{\rho_2}{\rho_1} + x\right) \frac{l}{\sqrt{C_{p1} T}}. \quad (2)$$

The total density

$$\rho = \left(\frac{1-x}{\rho_1} + \frac{x}{\rho_2}\right)^{-1}, \quad (3)$$

so if  $x > \rho_2/\rho_1$ , then the vapor occupies most of the volume. Notice that in the very low  $x$  limit  $c \propto \rho_2/\rho_1$ , whereas in a two-component gas-solid/liquid mixture with the solid/liquid providing most of the mass,  $c \approx c_2(\rho_2/\rho_1)^{1/2}$ . This demonstrates that a single-component medium is more compressible in the low- $x$  limit than a two-component medium.

It is useful to explain this difference in behavior in detail. For a single-component medium in phase equilibrium, the Clausius-Clapeyron equation implies

$$\Delta P = \frac{l}{T(1/\rho_2 - 1/\rho_1)} \Delta T \approx \frac{l p_2}{T} \Delta T. \quad (4)$$

When  $x \ll 1$ , almost all the internal heat is in phase 1, so the condition of adiabaticity is  $M_1 C_{p1} \Delta T \approx l \Delta M_1 = -l \Delta M_2$ . Conservation of mass implies  $\Delta M_2 = \rho_2 \Delta V_2 + V_2 \Delta \rho_2 \approx \rho_2 \Delta V_2$  since  $\Delta M_2/M_2 \propto M_1/M_2 \propto x^{-1}$ , whereas  $\Delta \rho_2/\rho_2 \propto x^0$ . Phase 1 is much less compressible than phase 2, so  $\Delta \rho \approx -\rho \Delta V_2/V$ . Combining these expressions we find that

$$\frac{\Delta T}{T} \approx \frac{l}{C_{p1} T} \frac{\rho_2}{\rho} \frac{\Delta \rho}{\rho}, \quad (4a)$$

and with equation (4) we recover equation (2). Equivalently,

$$\frac{\Delta T}{T} \approx -x \frac{l}{C_{p1} T} \frac{\Delta V}{V_2} < x \frac{l}{C_{p1} T}. \quad (5)$$

When  $x \ll 1$ , there is a limit to the wave amplitude if the system is not to transform to a single-phase medium. Large shifts in volume are accompanied by small shifts in temperature or gas density. It turns out (§ IV) that the protolunar disk has  $\rho_2/\rho_1 \sim 10^{-5}$ – $10^{-3}$ , so it can undergo substantial compression.

At temperatures  $kT/\mu \ll l$ , the two-phase sound speed reduces to the isothermal sound speed  $\sqrt{kT/\mu}$  in the limit  $x \rightarrow 1$  (vapor with a small mass fraction of liquid droplets). Equation (4) implies that for constant  $l$  the phase equilibrium condition takes the form

$$P = P_0 e^{-T_0/T}, \quad (6)$$

with  $T_0 = l\rho_2/P = l\mu/k$ . In the assumed temperature range we then have  $\Delta P/P \gg \Delta T/T$ , and the wave is isothermal to a good approximation. There is also a limit on the amplitude of waves in this case: the increase in the internal heat of the vapor upon adiabatic compression must be less than that required to vaporize the residual liquid.

The limited ability of the liquid, say, to conduct heat becomes important for wave frequencies  $\omega > \kappa/R^2$ . Here the length scale  $R$  is the characteristic size of the liquid drops when  $x$  is small, or the average distance between bubbles if the liquid percolates ( $x < \rho_2/\rho_1$ ). The quantity  $\kappa$  is the liquid thermal diffusivity. The effective liquid mass fraction is reduced for  $\omega$  in the above range. Similarly, the effects of the liquid surface tension  $\Sigma$  must be taken into account when the change in internal energy of the liquid during compression of the two-phase medium is comparable to the change in surface energy, i.e., when

$$R < \frac{2\Sigma}{C_{p1} T \rho_1}. \quad (7)$$

### III. GRAVITATIONAL INSTABILITY IN TWO-PHASE DISKS

In this section, we review some relevant facts about the stability of disks and consider the behavior of a gravitationally unstable two-phase disk. We limit ourselves to the case where the disk orbits a central body whose mass is much greater than the mass of the disk. In this case, the disk rotates differentially. A two-phase disk may settle into a state that is slightly gravitationally unstable. These slow instabilities continually heat the disk, and induce turbulence that can redistribute angular momentum.

#### a) Stability of a Thin Disk

The simplest way to test for stability is to consider linear axisymmetric perturbations of disk material. It is interesting to note that in the simpler case of viscous Couette flow, Taylor (1923) was remarkably successful in predicting instability by calculating the effects of axisymmetric perturbations. Nevertheless, it still has not been demonstrated that Couette flow is stable to two-dimensional nonaxisymmetric perturbations in the regime where it is stable to axisymmetric perturbations. (See, for example, Drazin and Reid 1981.) Grinfeld (1984) has studied the nonlinear stability of planar flow and obtained a partial answer to this problem. In the case of two-dimensional Couette flow, his results imply stability if there is a (rotating) frame in which the velocity is subsonic everywhere in the flow (and if, of course, the specific angular momentum of the flow increases outward).

In a similar fashion, the stability of disks to nonaxisymmetric perturbations is not fully understood. It is here that one must make an important distinction between Couette flow and disks. At least in the absence of viscosity, Couette flows that are stable to radial perturbations are the minimum energy configurations for axisymmetric flows with the given boundary conditions. The effect of viscosity, as Taylor showed, is only to suppress some instabilities.<sup>1</sup> But for orbiting disks, a lower energy configuration of disk material that conserves angular momentum can always be found by moving material inward in the inner part of the disk and outward in the outer part of the disk. The binding energy released in the inner parts is greater than that absorbed in the outer parts. Therefore, there is no firm reason to expect that disks are absolutely stable when they are stable to radial perturbations. The bottom line is that we have no direct experimental evidence on the stability of disks, except perhaps in the case of Saturn's rings.

<sup>1</sup> When the self-gravity of the fluid is taken into account, viscosity can actually induce slow, diffusive instabilities: e.g., the segmentation of a disk into rings. Clearly, this effect is irrelevant to the Taylor problem.

Recently, several authors have investigated the stability of both thick and thin disks (in which the effects of disk self-gravity can be neglected) to nonaxisymmetric perturbations. (See Papaloizou and Pringle 1984, 1985; Goldreich and Narayan 1985; Goldreich, Goodman, and Narayan 1986; Goodman, Narayan, and Goldreich 1987.) Thick, differentially rotating disks suffer from violent, global instabilities. Specifically, instabilities whose growth time is comparable to the orbital time exist in disks with power-law angular velocity profiles  $\Omega(r) \propto r^{-\alpha}$ , if the exponent  $\alpha$  is in the range  $3^{1/2} < \alpha < 2$ —that is, if the disk is thick. The instabilities in thin disks are much weaker. They depend on the existence of differential rotation and at least one boundary that can reflect the spiral density waves that redistribute energy within the disk. The growth time is therefore somewhat longer than the sound-crossing time of the disk, i.e., much greater than the orbital time.

To a first approximation, one investigates the stability of thin disks (vertical scale height  $h \ll r$ ) by treating them as two-dimensional sheets, with surface mass density  $\sigma$ . Then the local dispersion relation for radial perturbations proportional to  $e^{i(kr - \omega t)}$ , with  $kr \gg 1$  and  $kh \ll 1$  reads (Toomre 1964)

$$\omega^2 = k^2 \tilde{c}^2 - 2\pi G \sigma k + \kappa^2, \quad (8)$$

where the epicyclic frequency  $\kappa^2 = r^{-3}(d/dr)(r^4 \Omega^2)$  and  $\Omega$  is the local angular velocity. For Keplerian circular motion,  $\kappa = \Omega$ . Equation (8) has been derived in the approximation that the perturbations are independent of distance from the equatorial plane, so that the effects of pressure enter in a vertically averaged sound speed

$$\tilde{c}^2(h) = \frac{1}{\sigma(h)} \int_{-h}^{+h} \rho(z) c^2(z) dz, \quad (9)$$

where the surface mass density

$$\sigma(h) = \int_{-h}^{+h} \rho(z) dz. \quad (10)$$

A simple physical interpretation of the dispersion relation (8) is that the total radial restoring force is the sum of those due to pressure, gravity, and rotation. Short-wavelength perturbations are suppressed by pressure, and long-wavelength perturbations by the conserved angular momentum (see Fig. 1).

As discussed in § IV, the scale height must be defined carefully for a *two-phase* disk, in which the vapor phase is much less dense than the second phase. One must distinguish between the height of the photosphere,  $h_p$ , determined by the opacity of vapor and grains, and the average displacement of disk material from the equatorial plane,  $h_{\text{mass}}$  (see eq. [28]). In general,  $h_{\text{mass}}$  indicates the vertical distribution of the denser phase. For two-phase disks, it is  $\tilde{c}(h_{\text{mass}})$  that enters the dispersion relation, not  $\tilde{c}(h_p)$ .

In a pressureless, self-gravitating disk, instabilities grow progressively faster at shorter wavelengths; as first emphasized by Goldreich and Ward (1973), there is no characteristic length scale at which instabilities progress fastest unless the stabilizing effects of *both pressure and rotation* are taken into account.

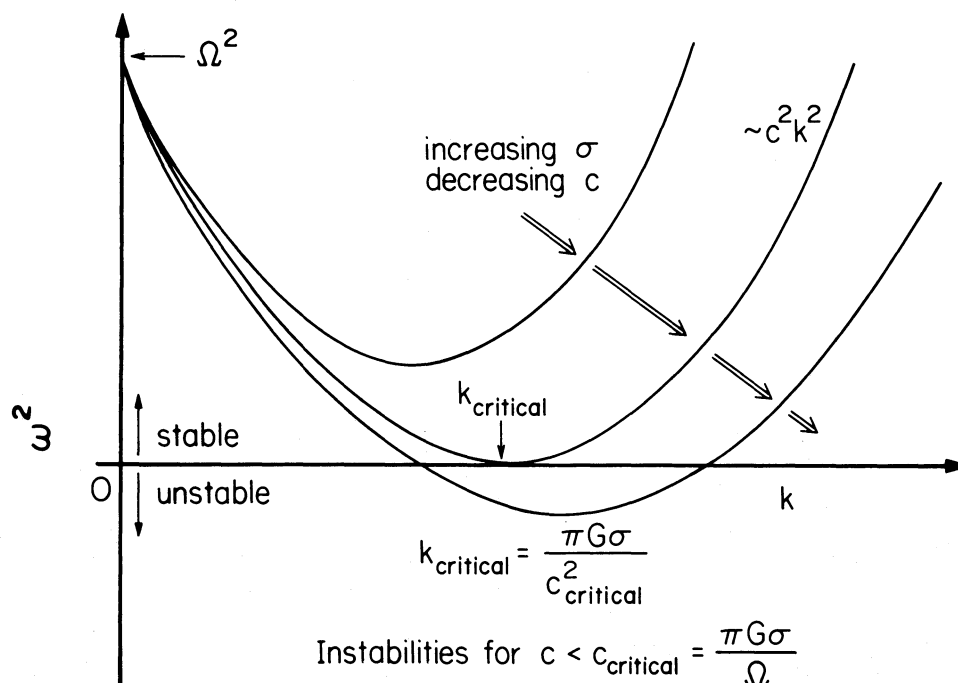


FIG. 1.—Dispersion relation of a thin disk under radial perturbations (eq. [8]), illustrating how a decrease in sound speed (at fixed surface density) can cause instability.



When the right-hand-side of equation (8) is negative for some  $k$ , the disk suffers from growing modes and is unstable. This occurs for

$$\tilde{c} < \tilde{c}_{\text{crit}} = \frac{\pi G \sigma}{\kappa}, \quad (11)$$

or, defining the stability parameter  $\eta \equiv \pi G \sigma / \kappa \tilde{c}$ , for  $\eta > 1$ . The fastest growing mode is

$$\kappa_{\text{crit}} = \frac{\pi G \sigma}{\tilde{c}^2}. \quad (12)$$

Once a disk is unstable by the above criterion, both axisymmetric modes, and nonaxisymmetric modes of comparable wavelength, should grow at comparable rates. As Goldreich and Ward (1973) have argued, the disk should break up in patches. (The essential ideas of this patch instability were first noticed by Edgeworth 1949.) When  $\eta = 1$ , the characteristic patch size is

$$l_{\text{crit}} = \frac{\pi}{k_{\text{crit}}} = \frac{\pi^2 G \sigma}{\kappa^2} \sim \frac{\pi}{2} h, \quad (13)$$

and the corresponding patch mass is

$$M_{\text{crit}} = \frac{\pi^2 G^2 \sigma^3}{\kappa^4} \sim \frac{\pi^3}{4} h^2 \sigma. \quad (14)$$

Most probably, a proper account of nonaxisymmetric effects would shift slightly the critical sound speed as derived above. We doubt that such an account would change the nature of the instability, i.e., drastically increase the critical sound speed. This is because the instabilities found above are *local*: they are insensitive to global properties such as the disk size, or conditions at its boundaries. On the other hand, the instabilities recently found in thin disks by considering the effects of nonaxisymmetric perturbations are very sensitive to boundary conditions. Their growth time is somewhat greater than the disk sound-crossing time, but local instabilities can grow much faster than this. The introduction of viscosity damps instabilities in Couette flows that are close to the boundary between stable and unstable regimes (constant angular momentum flows); it is possible that the same could happen in the case of self-gravitating disks.

In a marginally unstable single-phase disk,  $h_p$  is somewhat greater than twice  $h_{\text{mass}}$ , and in order of magnitude,  $h_p \sim h_{\text{mass}} \sim c/\Omega$ . In contrast, two-phase disks are much more centrally condensed:  $h_p \gg h_{\text{mass}}$ . In our application, we will find that  $h_{\text{mass}} \approx \tilde{c}_{\text{crit}}/\Omega$ . (See § IV for more details.) We conclude that  $k_{\text{crit}} h_p \sim \pi G \sigma / \Omega c \sim 1$  in a slightly unstable gaseous disk, and correspondingly  $k_{\text{crit}} h_{\text{mass}} \sim 1$  in a slightly unstable two-phase disk. In both cases, the thin disk condition  $k_{\text{crit}} h \ll 1$  is only marginally satisfied. Finally, an unstable gaseous disk is basically self-gravitating: the ratio of the vertical acceleration due to the disk matter, to that due to the central body is  $\sim 2\pi G \sigma(z)/\Omega^2 z > 1$ . The dense central layer of a two-phase disk may be self-gravitating, while its upper rarefied atmosphere is confined by the central potential. Typically  $k_{\text{crit}} h_p \gg 1$ , so that, in a sense, only the central layer is unstable.

In general, a thin gaseous disk has a cooling time much shorter than its spreading time  $t_{\text{visc}}$ , assuming that it instantaneously radiates the energy dissipated viscously:

$$t_{\text{cool}} \sim \frac{c^2 \sigma}{2 \sigma_{\text{SB}} T_p^4} \sim \frac{c^2}{G M_c u_r / r^2} \sim \left( \frac{h}{r} \right)^2 t_{\text{visc}}, \quad (15)$$

since  $h \sim c/\Omega$ . Here  $u_r$  is the radial velocity of the disk and  $M_c$  is the central mass around which it orbits. But in a two-phase disk

$$t_{\text{cool}} \sim \left( \frac{c_2}{\Omega r} \right)^2 t_{\text{visc}}, \quad (16)$$

where we have approximated the internal heat per mass of the disk by the vapor sound speed squared. That is,  $t_{\text{cool}}$  can be comparable to  $t_{\text{visc}}$  in a *thin* two-phase disk ( $h_{\text{mass}} \ll r$ )—the disk can have an extended optically thick atmosphere with  $c_2 \sim \Omega r$ , the orbital velocity.

### b) The Metastable State

Now we consider the behavior of a hot (vapor-fluid) disk that is gravitationally unstable by criterion (11). In particular, we argue for the existence of a *metastable* state, in which the disk hovers just over the edge of instability but does not break up in patches. (Paczynski 1978 treated a similar situation in the case of accretion onto a black hole.) The basic idea is that turbulence is generated by the abortive instabilities continually occurring in the disk. This turbulence induces an effective viscosity which causes the disk to spread, and at the same time heats the disk, maintaining the marginally unstable state. We will develop an explicit criterion for this mechanism to work.

It is clear that within the Roche radius (as defined for small, isolated liquid satellites) any condensations in a gravitationally unstable disk are tidally disrupted.<sup>2</sup> One expects a state intermediate between a smooth, continuous disk and a discrete set of satellites. Cameron and Ward (1978) first made this point for the case of a protolunar disk in Earth orbit.

If a disk is unstable, however, one must ask whether it may self-consistently radiate the heat generated viscously within it.

<sup>2</sup> In this discussion, by "Roche radius" we mean the Roche radius as evaluated for a satellite with the density of the fluid phase. For a silicate disk in Earth orbit,  $\rho \approx 3 \text{ g cm}^{-3}$  and  $r_{\text{Roche}} \approx 3 r_{\oplus}$ .

Cameron and Ward (1978) assumed that any ejecta from a giant impact would cool and solidify in Earth orbit, but their hypothetical particulate disk spreads so quickly (in  $\sim 0.1$  yr) that, radiating as a blackbody, it can dispose of only a tiny fraction of the binding energy released. Indeed, Cameron and Ward assumed that the disk instabilities induced turbulence only within the Roche radius, and that the disk fragmented exterior to this. Instead, we conclude that the disk can remain intact and marginally unstable *outside* the radius, so long as the growth time of local instabilities is longer than the orbital time.

There are two conditions for the existence of this metastable state:

1. The disk must enter the unstable regime from a stable state. This happens in the case of the giant impact hypothesis for the origin of the Moon, if the disk forms from a cloud of debris. As the liquid component settles toward the midplane, which is perpendicular to the cloud's net angular momentum, the two-phase sound speed decreases as the vapor mass fraction decreases, and eventually the disk becomes gravitationally unstable. Material may settle directly outside the Roche radius, as some of the recent simulations of Benz, Slattery, and Cameron (1987) suggest.

2. The gravitational binding energy of the disk must be large enough that it spreads very slowly on an orbital time scale. As we argue below, the effective viscosity due to the disk turbulence (and therefore the spreading rate) increases as the disk becomes more unstable. A long spreading time implies marginal instability.

There are many conceivable ways in which photospheric losses may saturate. For example, if the photospheric and central temperature of the disk are too high, the condition for instability may be lost. The sound speed rises rapidly in a single-component, two-phase medium as one approaches the critical temperature. It may be, as is the case for the protolunar disk, that the sound speed at the critical point is too high for instability.

However, it is not this mechanism that causes the instabilities to saturate in the protolunar disk, but one peculiar to two-phase disks. The equation of state for a two-phase medium is particularly simple:  $P = P(T) = P_0 e^{T_0/T}$ . Now the photospheric pressure is  $P_p = g/\kappa$  in a plane-parallel atmosphere in local thermodynamic equilibrium. The height of the photosphere is

$$h_p \sim \Omega^{-1} \sqrt{\frac{kT_p}{\mu}}, \quad (17)$$

since the range of temperature within the disk between mid-plane and photosphere is typically no more than a factor of 2, being buffered by the exponential dependence of pressure on temperature in a two-phase medium. But  $T_p = T_0 [\ln(P_0/P_p)]^{-1}$  and  $P_p = \Omega^2 h_p / \kappa$ , yielding, together with equation (17),

$$h_p^2 = \frac{kT_0}{\Omega^2 \mu} \left[ \ln \left( \frac{P_0 \kappa}{\Omega^2 h_p} \right) \right]^{-1}. \quad (18)$$

We may determine  $h_p$  and  $T_p$  once we know the constants  $P_0$ ,  $T_0$ , and  $\kappa$ . Neglecting the logarithm,  $h_p$  scales with radius as  $r^{3/2}$ .

We will find in § IV that the equilibrium state of a protolunar disk is very hot and puffy, with the liquid phase occupying only a small fraction of the volume, even in the mid-plane region. One might suspect that the liquid component would settle to the mid-plane, substantially decreasing the sound speed and leading to the breakup of the disk. Now, as settling occurs and the disk becomes more unstable, the local dissipation increases. In order to determine whether or not settling is prevented, we must compare the energy dissipated in one settling time, with the vertical potential energy liberated.

The vertical equation of hydrostatic equilibrium in the disk is  $\Omega^2 z + 2\pi G \sigma(z) = -(1/\rho)(\partial P/\partial z) = -(x/\rho_2)(\partial P/\partial z)$ . Since the two terms on the left are comparable in a marginally unstable disk, the potential energy of material above the mid-plane is  $\sim \Omega^2 z^2 \sim x(kT/\mu)$  per unit mass. By the time the disk has spread substantially in the radial direction, it has liberated orbital potential energy of order  $\Omega^2 r^2$  per unit mass. Therefore, the instabilities occurring in the disk may heat it sufficiently to prevent sedimentation if

$$\frac{t_{\text{sed}}}{t_{\text{visc}}} > x \frac{kT}{\mu \Omega^2 r^2}. \quad (19)$$

Here  $t_{\text{sed}}$  is the sedimentation time, and  $t_{\text{visc}}$  is the viscous spreading time. In the next section, we will consider this question in detail for the protolunar disk.

This argument immediately implies that the mass scale height  $h_{\text{mass}}^2 \sim x(0)(kT_p/\mu\Omega^2)$ , since the temperature varies slowly through the disk. That is, there is a relation between  $h_{\text{mass}}$  and  $h_p$ :

$$h_{\text{mass}} \sim \sqrt{x(0)h_p}. \quad (20)$$

Here  $x(0)$  is the vapor mass fraction at the mid-plane.

It should be noted that the above mechanism limits *any* local instability in the disk. This contrasts with the situation studied by Paczyński (1978), namely a (single-phase) disk fed by a constant external source. In such a case, the high surface density required for gravitational instability implies a low viscosity and marginal instability, but in fact the surface density is not a predetermined quantity. Paczyński assumes that if the surface density decreases and the disk becomes stable, there is no other source of viscosity. If this is not so, then a high surface density, marginally unstable disk is not the preferred state.

The ultimate source of the energy dissipated viscously in the disk is the disk's orbital binding energy.<sup>3</sup> When a patch of the disk undergoes instability, a significant fraction of its self-binding energy is transformed into turbulent motion, and thence into heat. But when the disk is only marginally unstable, the overdense patch is disrupted by the disk shear, as we now demonstrate. The tidal

<sup>3</sup> The latent heat released when vapor condenses onto the liquid or solid phase must also be radiated by the disk. As is discussed in § IV, this may dominate the viscous dissipation in the outer part of a protolunar disk.

acceleration at the edge of the condensation is  $a_{\text{tid}} \sim \Omega^2 l_{\text{crit}}/2$ . The gravitational acceleration parallel to the disk due to the (localized) overdensity is  $a_{\text{sg}} \sim G \Delta M / (l_{\text{crit}}/2)^2 \sim 4G \Delta \sigma$ . From equation (13),  $l_{\text{crit}} \sim \pi^2 G \sigma / \Omega^2$  in a marginally unstable disk, and

$$\frac{a_{\text{sg}}}{a_{\text{tid}}} \sim \frac{\Delta \sigma}{\sigma}. \quad (21)$$

If  $\eta - 1 \ll 1$ , then the instability grows in a time  $2\pi[\omega(k_{\text{crit}})]^{-1} \approx 2\pi\Omega^{-1}(\eta - 1)^{-1} \gg 2\pi\Omega^{-1}$ , and the local increase in the surface density in one orbit is  $\Delta\sigma \sim 4\sigma(\eta - 1) \ll \sigma$ . So, even outside the Roche radius, gravitationally induced clumps are shorn apart by differential rotation if the disk is marginally unstable.

Should a part of the disk be so unstable that local instabilities grow in a time short compared to the orbital period, it will break up into lumps of mass  $\sim M_{\text{crit}}$  (eq. [14]). The density of these lumps will be comparable to that of the solid or liquid component, and they will be stable outside the corresponding Roche radius:

$$r > r_{\text{Roche}} = 2.455 r_c \left( \frac{\rho_c}{\rho_1} \right)^{1/3}, \quad (22)$$

where  $\rho_c$  is the mean density of the central body and  $r_c$  is its radius.

We can now predict the behavior of a gravitationally unstable two-phase disk that is turbulently heated outside the Roche radius (that satisfies conditions 1 and 2). As it spreads,  $\sigma$  drops and the disk becomes more unstable in order to maintain the local, photospherically regulated dissipation. Eventually, instabilities in the outer parts of the disk grow on an orbital time scale, and the disk breaks up into separate lumps. The radius at which this occurs gradually moves inward, until it coincides with the Roche radius. Further in than this, the disk remains turbulent but lumpy, until the instabilities are no longer limited by the photospheric opacity, and the disk can cool.

One should note that the vigorous global instabilities discovered by Papaloizou and Pringle could occur in the protolunar disk if it forms by the settling of a vapor-fluid cloud, while the disk is still thick. It is not clear what redistribution of the disk material might result. Once the disk thins out, the role of global, nonaxisymmetric instabilities is limited. Since, in this model, the strength of instabilities is limited by the ability of the disk to radiate heat, it is possible that nonaxisymmetric instabilities could be responsible for the disk dissipation. This could happen if the required growth time of ordinary local instabilities were greater than the disk sound-crossing time. If stronger instabilities were required, however, they could only be local: the orbital period is generally shorter than the disk sound-crossing time. Our conclusion that the disk breaks up when the growth period of local instabilities becomes comparable to the orbital period should then not be modified.

It remains to estimate how the disk viscosity  $\nu$  depends on the stability parameter  $\eta$ . Before we do this, we must emphasize that the thermal state and evolution of the disk do not depend on this prescription, provided the self-regulated state described above is achievable. If the disk vertical structure is known so that the photospheric temperature may be calculated, then the sum of local dissipation and secular cooling is determined. As equation (16) reveals, these two effects may be comparable. In terms of other disk variables, the viscosity is

$$2\sigma_{\text{SB}} T_p^4 = \nu \sigma \left( r \frac{d\Omega}{dr} \right)^2 + \left. \frac{dE}{dt} \right|_{\text{cool}}. \quad (23)$$

The second term on the right side of equation (23) represents secular cooling. In § IV, we consider a specific model in which this term is dominated by the release of latent heat as the vapor phase condenses.

One may estimate the disk viscosity from the above stability considerations. For  $\eta > 1$  and  $\eta - 1 \ll 1$ , instabilities of characteristic size  $l_{\text{crit}}$  occur in the disk, with growth rate  $(\eta - 1)\Omega$ . The random velocities induced in the disk will then be of magnitude  $v_{\text{turb}} \sim (\eta - 1)\Omega l_{\text{crit}}$ , correlated over a length scale  $\sim l_{\text{crit}}$ . We expect the size of the turbulent eddies to be comparable to the distance that they travel in one orbital time, over which they are disrupted. Thus  $l_{\text{turb}} \sim (\eta - 1)l_{\text{crit}}$ , and the turbulent viscosity is

$$\nu_{\text{turb}} \sim l_{\text{turb}} v_{\text{turb}} \sim \Omega l_{\text{crit}}^2 (\eta - 1)^2 \sim \frac{(\pi^2 G \sigma)^2}{\Omega^3} (\eta - 1)^2. \quad (24)$$

#### IV. FORMATION OF THE MOON FROM A HOT CIRCUMTERRESTRIAL DISK

In this section, we work out the properties of a circumterrestrial disk of sufficient mass to form the Moon and construct simple models of its evolution. The disk is so massive, and therefore so dissipative, that it should have a hot, liquid-vapor composition. As discussed further below, we approximate the liquid as  $\text{SiO}_2$ ; the coexisting vapor is mainly a mixture of  $\text{SiO}$  and  $\text{O}_2$ . Such a disk is a possible product of the ejecta from an impact between the proto-Earth and a giant planetesimal. We will show below that while hot, the disk is subject to gravitational instabilities as described in § III, and that these instabilities maintain the disk at a photospherically regulated central temperature of  $\sim 3000$  K. The spreading time of the disk is longer than its secular cooling time.

Before proceeding, we remark that it is not clear whether a giant impact results in an orbiting disk. The numerical experiments of Benz, Slattery, and Cameron (1986) initially indicated that a fraction of the mass of the impactor would survive the impact as a single, gravitationally bound body. (Their simulations assumed a small velocity at infinity, but a total angular momentum equal to that of the present Earth-Moon system.) An extrapolation of the highly eccentric orbit of this body brought it within the Roche limit, where presumably it would disintegrate and spread into a disk. Such a disk-forming process is different from the one we initially envisaged, in which a large fraction of the shocked material from the impactor and impacted region of the Earth is vaporized (Stevenson 1985). Pressure-gradient forces within this two-phase cloud may sufficiently distort the orbits of cloud material



that it settles directly into a disk. It is possible that (Stevenson 1987) the smoothed-particle hydrodynamics code of Benz *et al.* did not correctly account for the expected two-phase composition of the shocked material and considerably underrepresented these pressure-gradient forces.

More recent simulations by Benz, Slattery, and Cameron (1987) have suggested another mechanism by which ejecta from the impact might reach Earth orbit: tidal torques. In these simulations, the projectile has a core. The core and mantle of the impactor separate after impact, with the core essentially intact in the particular case where the projectile mass is  $M_{\oplus}/7$ . The denser core is captured by the Earth, and some material from the impactor mantle reaches an orbit via a three-body interaction. In most of these simulations, material is placed in orbit both inside and outside the Roche radius, but in a few, all the material lies outside the Roche radius.

#### a) Early Evolution of the Disk

Consider the case where a cloud of ejecta in Earth orbit settles directly into a disk. We make note of two important inequalities. First, the cooling time of the cloud greatly exceeds the sedimentation time. Second, the sedimentation time greatly exceeds the orbital period. The first inequality means that the initial state of the disk is hot. The second inequality means that the ejecta are sheared out into a disk before the two-phase sound speed becomes so low that the ejecta are gravitationally unstable. (Recall that the two-phase sound speed  $c$  is proportional to the vapor mass fraction  $x$  (eq. [2]) in the limit  $1 \gg x \gg \rho_2/\rho_1$ . As the condensed phase settles toward the midplane of the disk, both  $x$  and  $c$  decline.)

The terminal velocity  $v_z$  of the liquid component toward the equatorial plane is given by  $a^2 \rho_2 v_z^2 \sim \rho_1 a^3 G M_{\oplus} z / r^3$  at high Reynolds number, implying a sedimentation time

$$t_{\text{sed}} \sim \frac{z}{v_z} \sim \frac{1}{2\pi} \sqrt{\frac{\rho_2}{\rho_1}} \frac{z}{a} t_{\text{orb}}. \quad (25a)$$

Here  $a$  is the characteristic size of the liquid drops. (The terminal velocity is reached in a time  $\ll t_{\text{sed}}$  as long as  $t_{\text{sed}} \gg t_{\text{orb}}$ .) But a liquid drop falling near terminal velocity will break up on a time scale  $\sim a/v_z$  unless held together by surface tension; we may therefore estimate the drop size as  $2\Sigma/a \sim \rho_2 v_z^2$ , or (the surface tension  $\Sigma$  is  $\sim 300$  ergs  $\text{cm}^{-2}$  for liquid silicates; Walker and Mullins 1981)

$$a \sim \sqrt{\frac{\Sigma r^3}{\rho_1 G M_{\oplus} z}} \sim 1 \text{ cm} \left( \frac{r}{r_{\oplus}} \right), \quad (25b)$$

where we assume that the initial height of most of the ejecta is  $z \sim r$ . Notice that the drop size does not depend on  $\rho_2$  (an uncertain quantity), but  $t_{\text{sed}}$  does.

We may estimate the sedimentation time, by assuming that a fraction  $\epsilon_2$  of the ejected mass  $M_d$  is vapor.<sup>4</sup> Then the density of vapor in the ejecta is  $\rho_2 r^3 \sim \epsilon_2 M_d$ , or  $\rho_2 \sim 0.3 \epsilon_2 (M_d/M_{\oplus})(r/r_{\oplus})^{-5}$ . The sedimentation time is

$$\frac{t_{\text{sed}}}{t_{\text{orb}}} \sim 10^3 \epsilon_2^{1/2} \left( \frac{M_d}{M_{\oplus}} \right)^{1/2} \left( \frac{r}{r_{\oplus}} \right)^{-3/2}. \quad (25c)$$

Finally, the cooling time of a disk of radius  $r_d$  and (uniform) temperature  $T$  is

$$t_{\text{cool}} \sim \frac{M_d C_p T}{2\pi r_d^2 \sigma_{\text{SB}} T^4} = 50 \text{ yr} \left( \frac{r_d}{r_{\oplus}} \right)^{-2} \left( \frac{T}{2000 \text{ K}} \right)^{-3} \left( \frac{M_d}{2M_{\oplus}} \right). \quad (26)$$

Thus, the cooling time is significantly longer than the orbital time.

#### b) Equation of State of a Two-Phase Silicate Disk

The calculation of the sound speed for liquid-vapor silicates is described in the Appendix. In Figure 2, we plot the sound speed at various temperatures, as a function of the vapor mass fraction  $x$ , for two different two-phase equations of state. Notice the large decrease in the sound speed at low  $x$ , by as much as six orders of magnitude from the vapor sound speed. There is considerable disagreement between the equation of state for  $\text{SiO}_2$  as measured by Ruff and Schmidt (1921) and that calculated by Krieger (1967). For example, the pressure at  $T = 2500$  K is 1 atm, and 0.01 atm, and the derived heat of vaporization  $1.2 \times 10^{10}$  and  $1.6 \times 10^{10}$  ergs  $\text{g}^{-1}$ , respectively. One should note that the use of the Ruff and Schmidt  $P$ - $T$  curve at temperatures above 1500 K is a considerable extrapolation. We employ the Krieger curve in all our calculations hereon. Similar numbers apply to more realistic silicate compositions (e.g.,  $\text{MgSiO}_3$ ), but we avoid analysis of these because of the problems of incongruent vaporization.

A good indicator of the difference between the two phase-equilibrium curves is the value of the vapor density, which we plot in Figure 3. The vapor density calculated from the Krieger curve is consistently one to two orders of magnitude smaller than that calculated from the Ruff and Schmidt curve, for temperatures in the range 2000–3500 K.

The differences in disk structure that result from using the two equations of state are very small, however, because the pressure in the disk is basically fixed by the surface density of the disk. The temperature  $T = T_0 / \ln(P_0/P)$  varies very slowly with  $P$ , and depends only weakly on the assumed values of  $T_0$  and  $P_0$ .

A nominal disk surface density is two Moon masses spread uniformly over a disk of outer radius  $3r_{\oplus}$ , namely  $1 \times 10^7$  g  $\text{cm}^{-2}$ . Assuming the mass of the proto-Earth to be the same as that of the present Earth, equation (11) implies

$$\tilde{c}_{\text{crit}} = 2400 \eta^{-1} \left( \frac{r}{r_{\oplus}} \right)^{3/2} \sigma_7 \text{ cm s}^{-1}. \quad (27)$$

<sup>4</sup> To be more specific, we take  $\epsilon_2$  to be the vapor mass fraction when the ejecta have been spread out into a disk, but before the disk has settled significantly in the vertical direction.



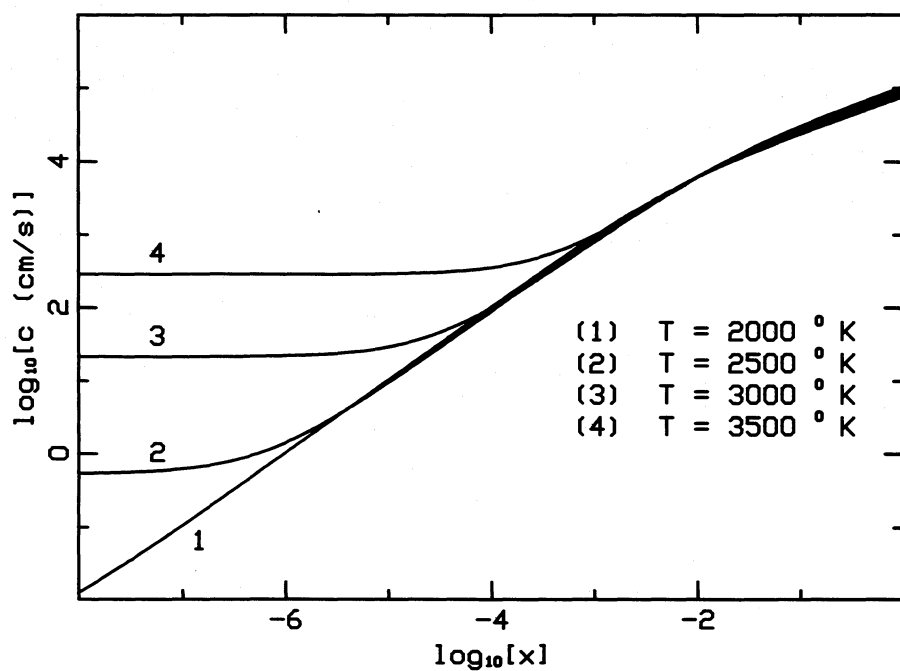


FIG. 2a

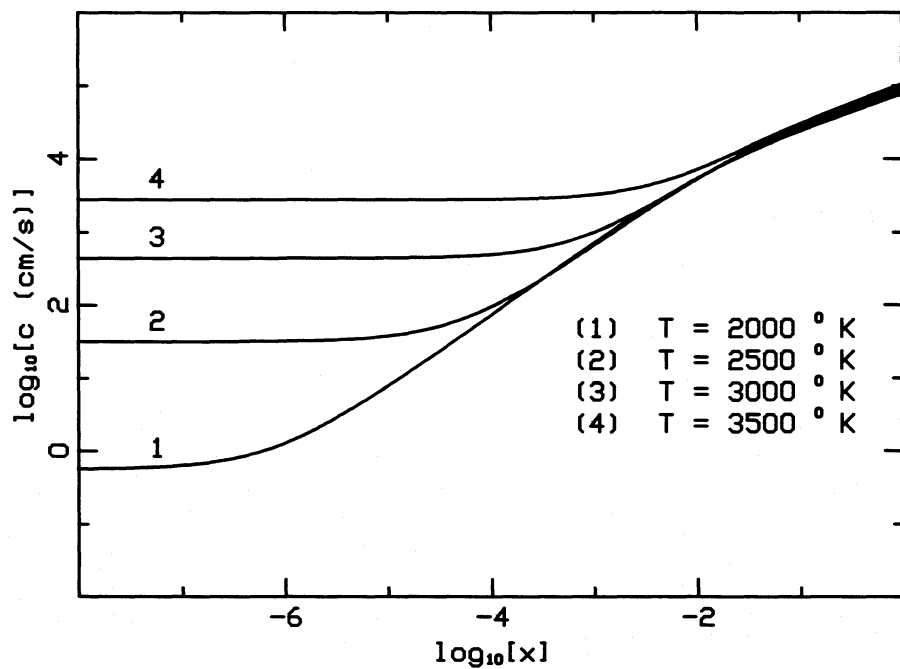


FIG. 2b

FIG. 2.—Sound speed  $c$  for a two-phase medium (liquid  $\text{SiO}_2$  and coexisting vapor) as a function of the vapor mass fraction  $x$ , using the phase equilibrium curve of (a) Krieger (1967), and (b) Ruff and Schmidt (1924). See text and Appendix for details of the computation.

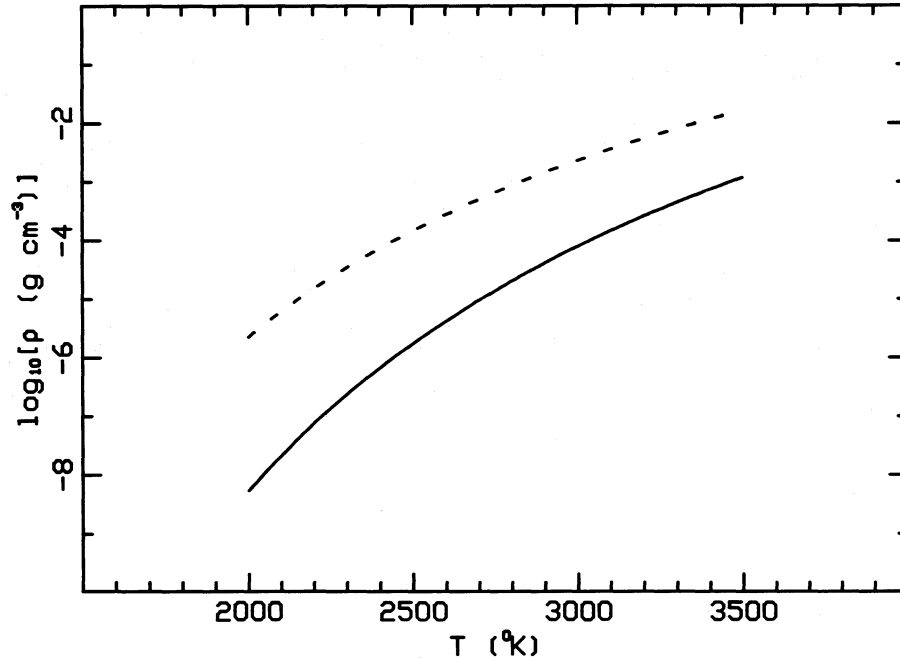


FIG. 3.—Density  $\rho_2$  of vapor  $\text{SiO}_2$  coexisting with the liquid phase, as a function of temperature  $T$ . Notice the large difference resulting from the use of the phase equilibrium relation of Krieger (1967) (solid curve) and that of Ruff and Schmidt (1924) (dashed curve).

Here  $\sigma = \sigma_7 \text{ g cm}^{-2}$ . Shortly, we will estimate the mid-plane temperature and vapor mass fraction of a marginally unstable disk. We will find that  $\rho_2/\rho_1 \ll x$  unless  $\sigma_7 \ll 1$ , so that the liquid occupies only a small fraction of the volume, even at the disk mid-plane.

### c) Vertical Structure of a Disk

We now consider the vertical structure of the disk. The dissipation rate is so high that the disk is convecting vigorously. We therefore expect it to be well-mixed, with the phase equilibrium condition for liquid-vapor silicates holding up to the photosphere. The Krieger (1967) calculation may be parameterized by a relation of the form (6), with  $P_0 = 3.2 \times 10^{14} \text{ dynes cm}^{-2}$  and  $T_0 = 6.0 \times 10^4 \text{ K}$ . The main uncertainty in the vertical structure lies in the vertical dependence of the vapor mass fraction,  $x$ . Equivalently, it is not clear whether the liquid component will “rain out.” If it does not, then the vertical structure is easily calculated. We consider the question of rainout later in this section.

In two-phase disks, it generally holds that the photospheric height  $h_p \gg \tilde{c}_{\text{crit}}/\Omega$ , and the mass-weighted scale height  $h_{\text{mass}} \approx \tilde{c}_{\text{crit}}/\Omega = \pi G \sigma / \Omega^2$ , where

$$h_{\text{mass}} \equiv \frac{2}{\sigma} \int_0^{h_p} z \rho(z) dz. \quad (28)$$

[When  $\rho(z) = \text{constant}$ ,  $h_{\text{mass}} = \frac{1}{2} h_p$ .] Notice that whereas  $h_p$  scales as  $r^{3/2}$  (eq. [18]),  $h_{\text{mass}}$  scales as  $r^3$ .

Our calculations include both the component of Earth's gravity perpendicular to the disk and the disk's self-gravity. In the range of  $\sigma$  of interest to us, the central layers of the disk are marginally self-gravitating. The vertical equation of hydrostatic equilibrium reads

$$-\frac{1}{\rho} \frac{\partial P}{\partial z} = 2\pi G \sigma(z) + \Omega^2 z, \quad (29)$$

which integrates to give

$$P(0) = \frac{1}{2} \pi G \sigma^2 + \frac{1}{2} \Omega^2 \sigma h_{\text{mass}}. \quad (30)$$

At small  $z$  the ratio of the vertical component of disk gravity to that of Earth gravity is  $4\pi G \rho(0)/\Omega^2 > 2\pi G \sigma / \Omega^2 h_{\text{mass}} \approx 1$  in a marginally unstable disk. In other words, we expect

$$\rho(0) > 1.8 \left( \frac{r}{r_{\oplus}} \right)^{-3} \text{ g cm}^{-3} \quad (31)$$

in our disk models. An isolated satellite of density  $\rho(0)$  is stable only if

$$\rho(0) > (2.455)^3 \rho_{\oplus} \left( \frac{r}{r_{\oplus}} \right)^{-3} = 80 \left( \frac{r}{r_{\oplus}} \right)^{-3} \text{ g cm}^{-3}. \quad (32)$$

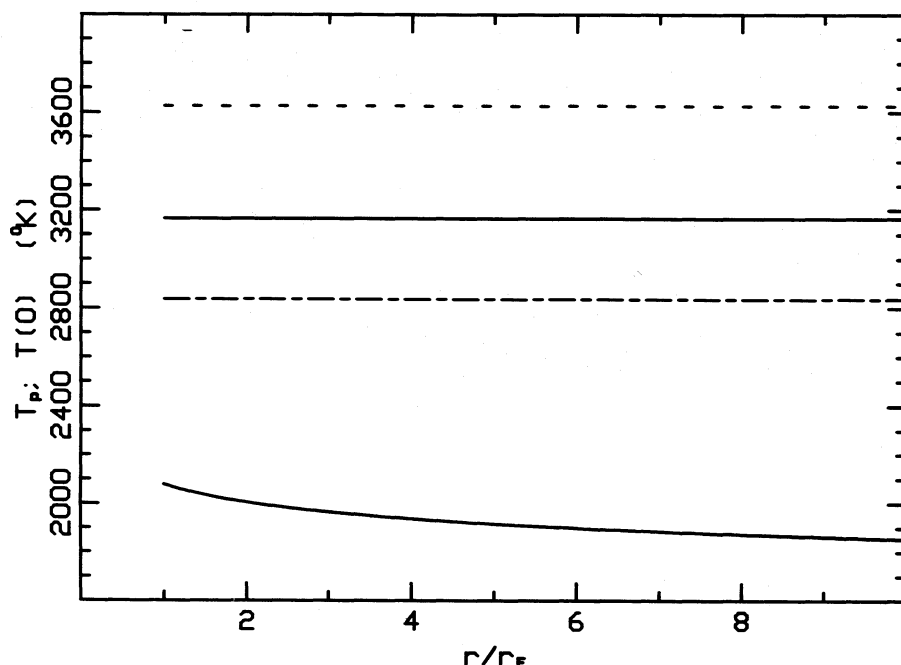


FIG. 4.—Radial dependence of photospheric temperature (lower curve) and mid-plane temperature (upper curves) of a massive silicate disk in Earth orbit. Markings correspond to different values of the surface density:  $10^7 \text{ g cm}^{-2}$  (dashed);  $3 \times 10^6 \text{ g cm}^{-2}$  (solid);  $10^6 \text{ g cm}^{-2}$  (double-dashed).

We find that the central density is never this high; so, in a sense, every part of the disk is within its Roche radius. In spite of this, our discussion of the fragmentation process in § III is still valid: when the instability growth time is less than the orbital time, an unstable patch condenses rapidly from  $\rho = \rho(0)$  to  $\rho \approx \rho_1$ , where  $\rho_1$  is the density of the condensed phase. If  $r > 3r_\oplus$ , this patch forms a stable proto-Moon.

The expression (30) for the mid-plane pressure of the disk allows us to obtain an estimate for the mid-plane temperature and vapor mass fraction at marginal stability ( $\eta = 1$ ). Since  $h_{\text{mass}} \approx \pi G \sigma / \Omega^2$ , the two terms in equation (30) are comparable and  $P(0) \approx \pi G \sigma^2$ . In other words, the contribution to  $P(0)$  from the disk's gravity and Earth's gravity are approximately equal. The temperature is then  $T(0) = T_0 / \ln [P_0 / P(0)]$ . The mid-plane temperature is plotted in Figure 4, as a function of distance from Earth, for several values of the surface density; it is in the range  $2500 < T(0) < 3500 \text{ K}$ . We conclude that a disk massive enough to form the Moon is necessarily very hot (too hot for its components to be solid).

Although  $P$  is function only of  $T$  in a two-phase medium, the density is not similarly determined. There is another independent thermodynamic variable, which we may take to be the vapor mass fraction  $x$ . If the disk enters the marginally unstable state under consideration, then the vertical mass-average of  $x^2$ ,

$$\bar{x}^2 = \frac{1}{\sigma} \int_{-h}^h x^2 \rho(z) dz, \quad (33)$$

is uniquely determined. To see this, let us first note that  $x$  satisfies the inequalities  $1 \gg x \gg \rho_2 / \rho_1$  inside the disk (we will justify this *a posteriori*). The two-phase sound speed (eq. [2]) is then  $c \approx x l (C_{p1} T)^{-1/2}$ . Accordingly, the vertical mass average of  $c^2$  (eq. [9]), which enters the dispersion relation (8), is proportional to the average of  $x^2$ . Since  $T$  varies slowly through the disk, we have

$$\bar{x}_{\text{crit}} \approx \bar{c}_{\text{crit}} \frac{\sqrt{C_{p1} T(0)}}{l} = \frac{\pi G \sigma}{\Omega l} \sqrt{C_{p1} T(0)} = 6.7 \times 10^{-4} \sigma_7 \left( \frac{r}{r_\oplus} \right)^{3/2} \left[ \frac{T(0)}{3000 \text{ K}} \right]^{1/2}. \quad (34)$$

In Figure 5, we plot  $\bar{x}_{\text{crit}}$  as a function of distance from Earth. In all cases,  $\bar{x}_{\text{crit}} \ll 1$ . For comparison, we plot the mid-plane vapor density  $\rho_2(0)$ . Notice that the liquid consistently occupies only a small fraction of the volume, since  $\bar{x}_{\text{crit}} \gg \rho_2(0) / \rho_1$  (where the density of the liquid phase  $\rho_1 \approx 3 \text{ g cm}^{-3}$ ).

When the disk is formed, it relaxes to vertical hydrostatic equilibrium. Initially,  $\bar{x}$  does not assume the value required for marginal instability. Indeed, a study of the thermodynamics of the giant impact which is assumed to have produced the disk (Stevenson 1985) indicates that the vapor mass fraction could initially be far larger than  $\bar{x}_{\text{crit}} \sim 10^{-3} - 10^{-2}$ , provided that the impact velocity  $\gtrsim 10 \text{ km s}^{-1}$ . The latent heat released as  $\bar{x}$  decreases to  $\bar{x}_{\text{crit}}$  is considerable. For example, the latent heat  $l = 1.7 \times 10^{11} \text{ ergs g}^{-1}$  (for  $\text{SiO}_2$ ) is equal to the gravitational binding energy to the Earth at a distance  $r = 4r_\oplus$ . Now, the presence of this excess vapor should not prevent the onset of gravitational instability as the condensed phase settles to the mid-plane of the disk. The young disk will have an extended vapor atmosphere, and the early photospheric losses may be dominated by the release of latent heat, rather than by viscous dissipation. We will see that the cooling time given by equation (26) is short compared to the spreading time of the disk.



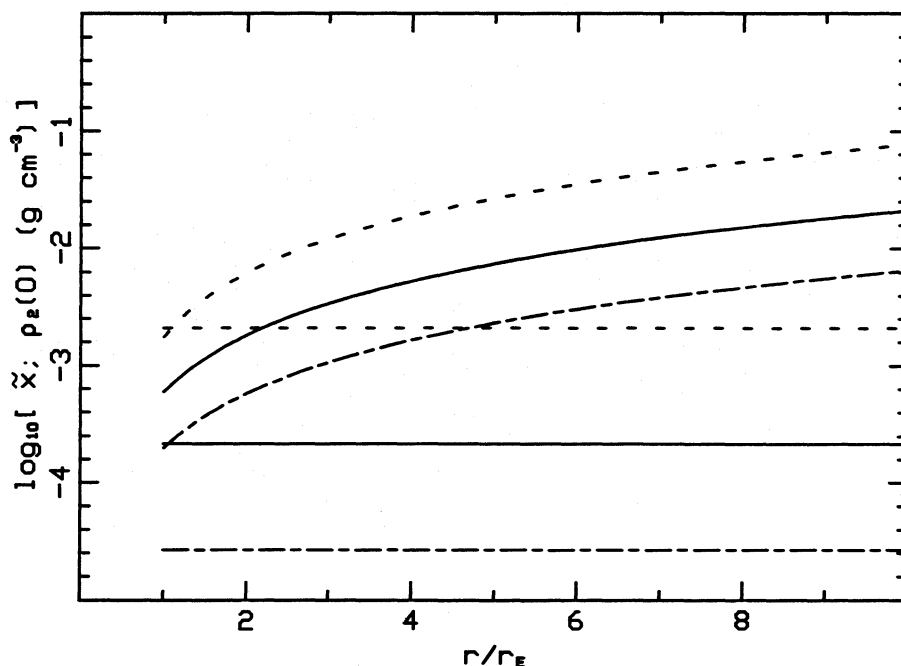


FIG. 5.—Radial dependence of vertically averaged vapor mass fraction  $\tilde{x}$  (top curves) and mid-plane vapor density  $\rho_2(0)$  (lower curves) for a marginally unstable silicate disk in Earth orbit. Markings correspond to different values of the surface density:  $10^7 \text{ g cm}^{-2}$  (dashed);  $3 \times 10^6 \text{ g cm}^{-2}$  (solid);  $10^6 \text{ g cm}^{-2}$  (double-dashed).

Of course, the mid-plane sound speed has a minimum value for fixed  $\sigma$ , which is reached when  $x < \rho_2(0)/\rho_1$ . This minimum value is

$$c_{\min}(0) = \frac{\rho_2(0)}{\rho_1} \frac{l}{\sqrt{C_{p1} T(0)}}. \quad (35)$$

Neglecting the logarithmic dependence of  $T(0)$  on  $\sigma$ ,  $\rho_2(0) \propto P(0) \propto \sigma^2$ , whereas  $\tilde{c}_{\text{crit}} \propto \sigma$  (eq. [11]). This means that  $c_{\min}(0)$  increases faster with  $\sigma$  than  $\tilde{c}_{\text{crit}}$ , and there is a maximum surface density (Fig. 6), above which a two-phase silicate disk in Earth orbit cannot undergo gravitational instabilities. Shortly, we will find a lower bound to  $\sigma$ .

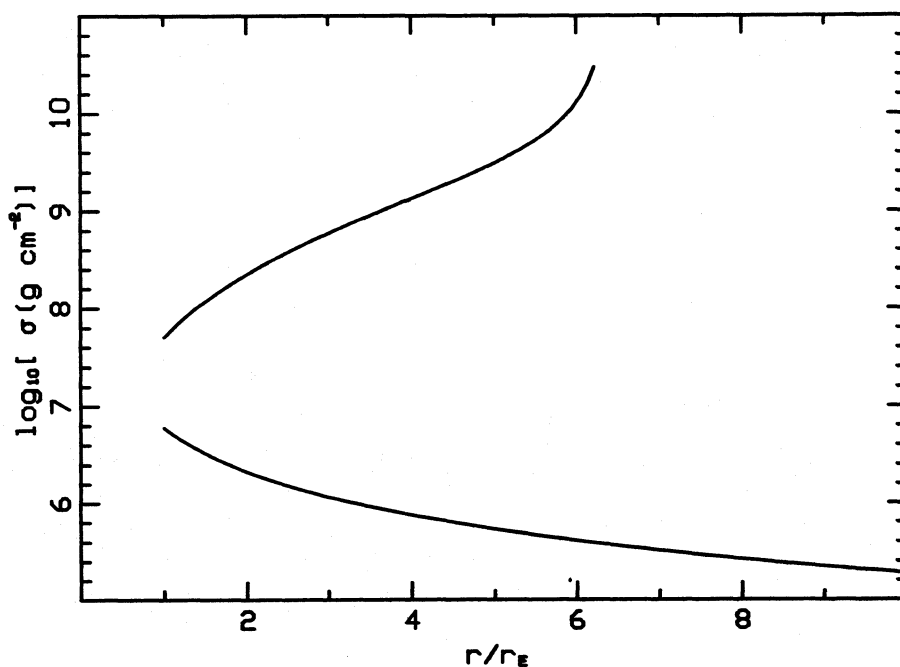


FIG. 6.—Maximum surface density of a hot silicate disk in Earth orbit that is gravitationally unstable (top curve); minimum surface density of such a disk that also has an adiabatic vertical profile (bottom curve).

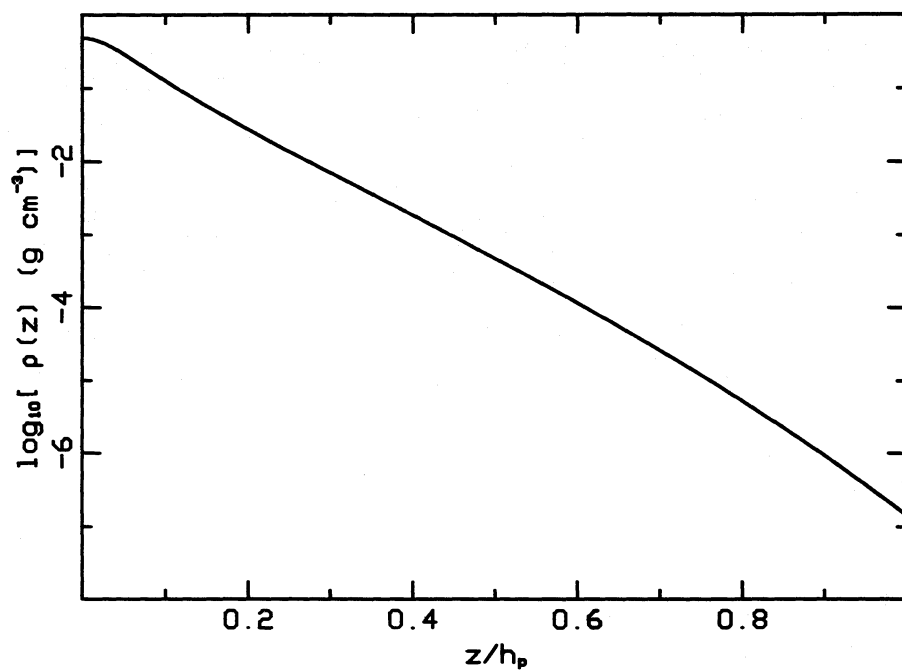


FIG. 7c

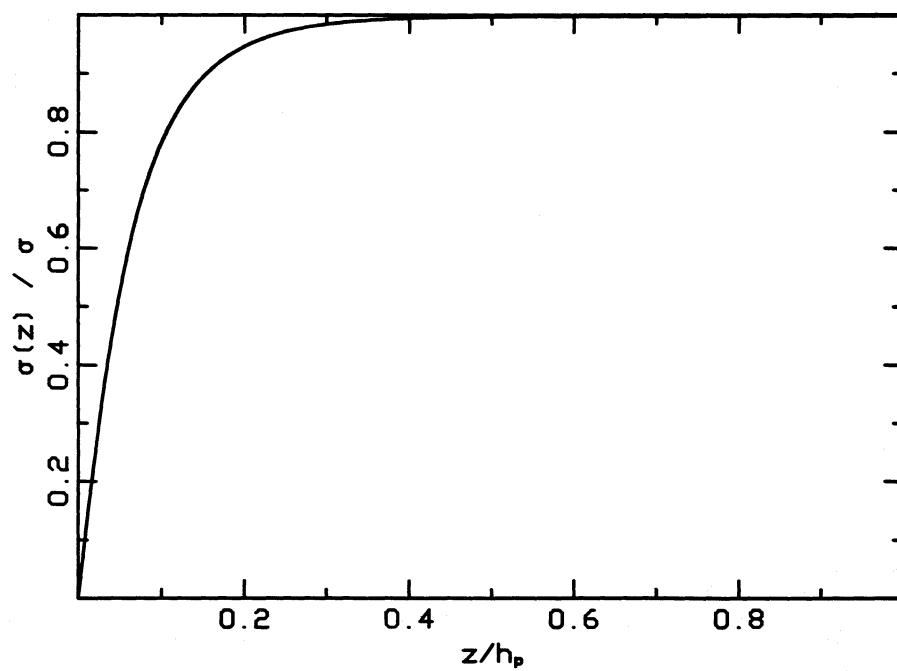


FIG. 7d

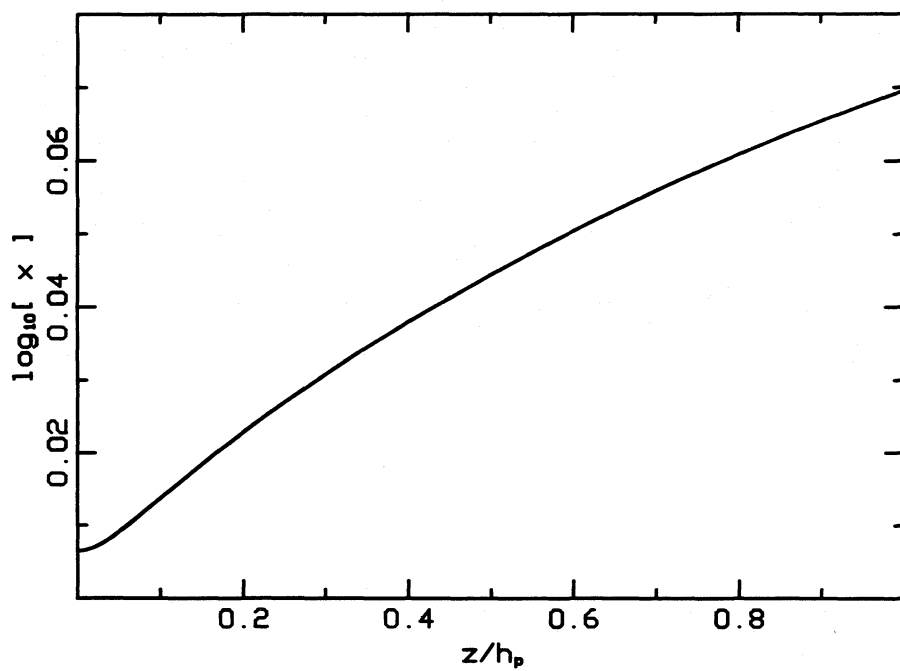


FIG. 7e

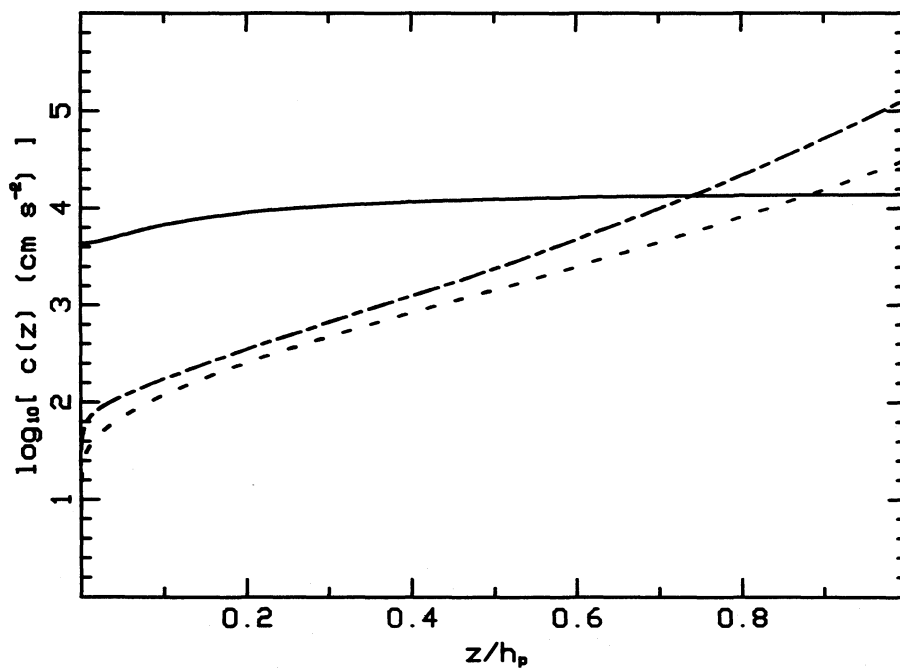


FIG. 7f



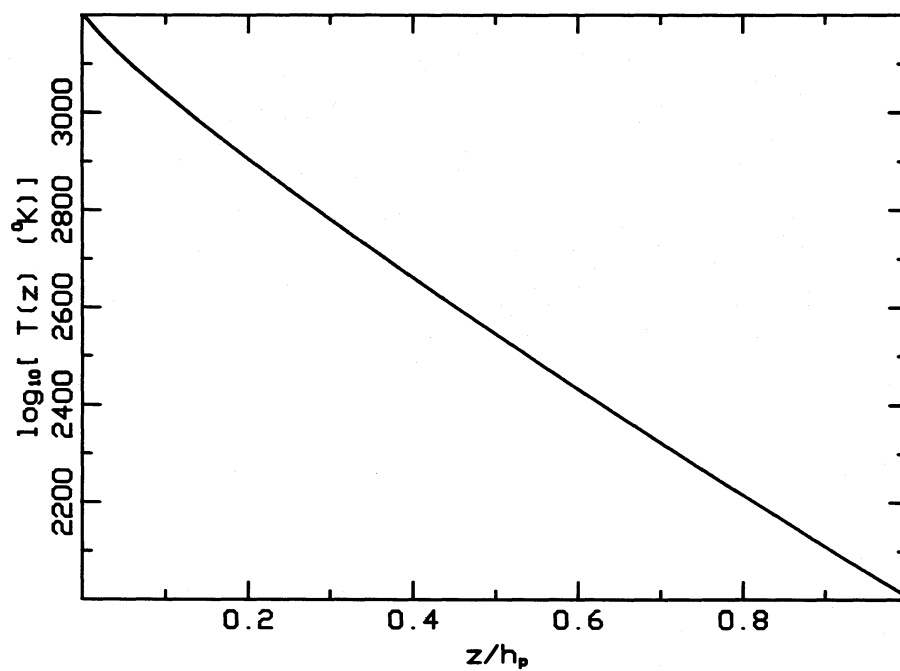


FIG. 8a

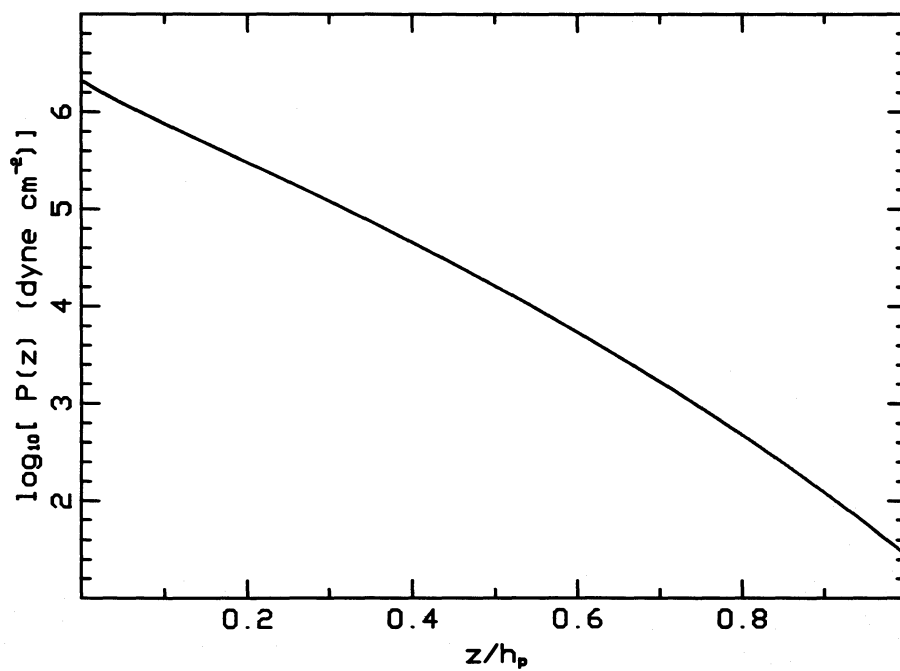


FIG. 8b

FIG. 8.—Variation of disk properties with height above mid-plane in units of photospheric height  $h_p$  (see Table 1). Here  $\sigma = 3 \times 10^6 \text{ g cm}^{-2}$  and  $r = 2r_\oplus$ . (a) Temperature; (b) pressure; (c) density; (d) cumulative surface density; (e) vapor mass fraction  $x$ ; (f) sound speed (*solid curve*, eq. [2]); convective velocity (*dashed curve*, eq. [51]); maximum rainout velocity (*double-dashed curve*, eq. [55]). Notice how  $x$  increases rapidly from the mid-plane.

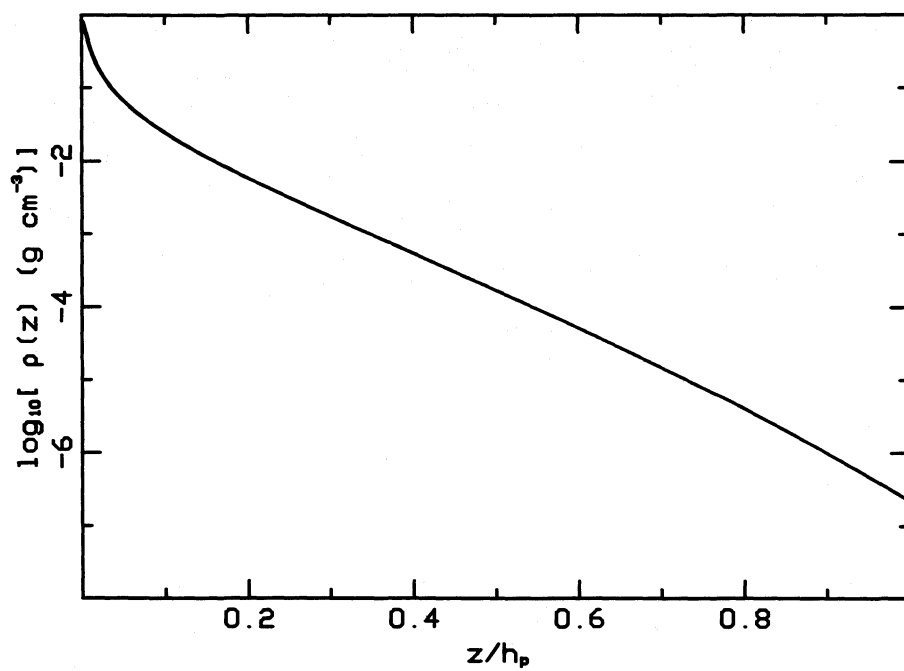


FIG. 8c

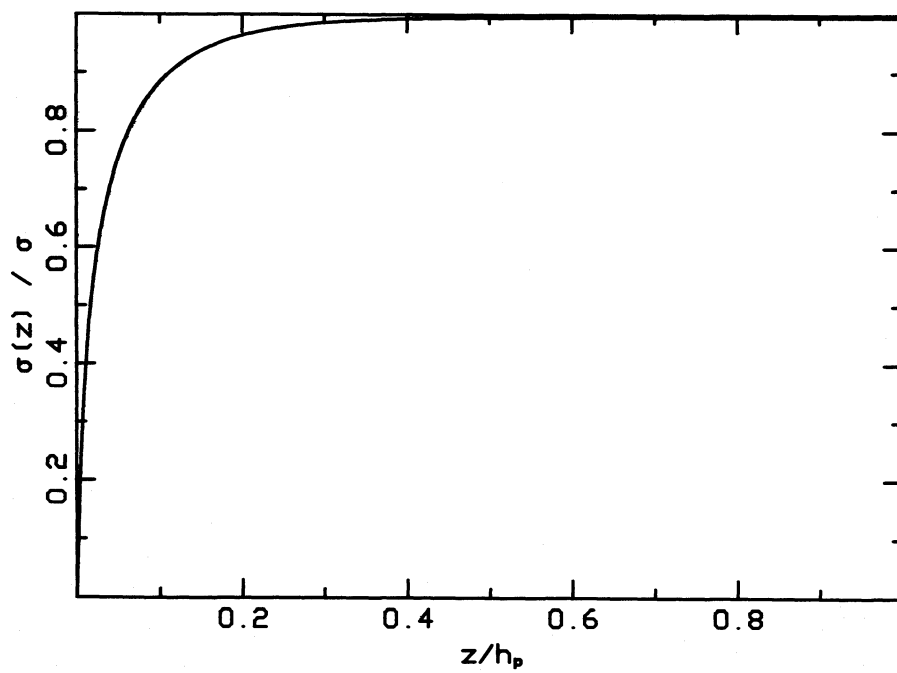


FIG. 8d

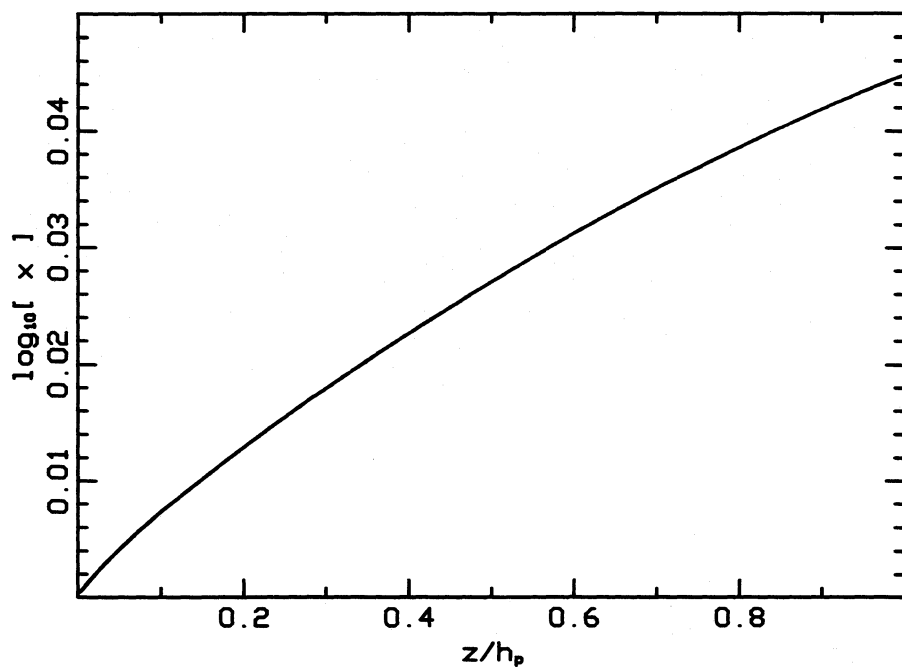


FIG. 8e

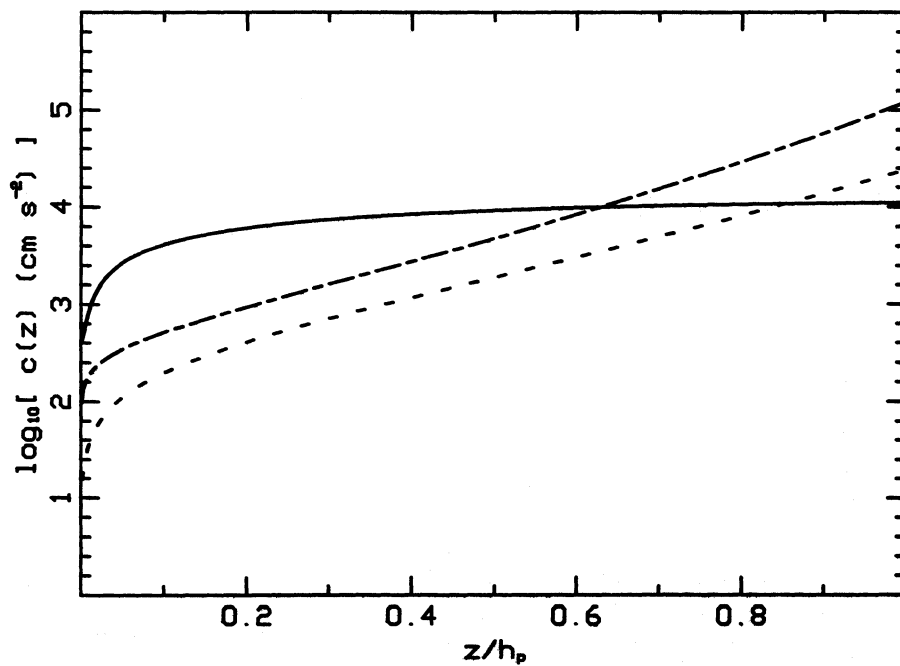


FIG. 8f



pressure  $P(0)$ , and integrating upward, simultaneously calculating  $\tilde{c}(h_{\text{mass}})$ . This value is compared with that needed for marginal stability, and the process repeated until the two values converge. We place the photosphere at a height  $h_p$  such that

$$\int_{-h_p}^{h_p} \rho(z) dz = \sigma,$$

and also require that  $P(h_p)$  satisfies  $P(h_p) \ll P(0)$ . No attempt is made to match these internal disk solutions to photospheric solutions.

Notice how the mass is centrally concentrated. Typically,  $h_{\text{mass}} \sim (1/10)h_p$ . The relation  $\tilde{c}_{\text{crit}} = h_{\text{mass}} \Omega$  holds to within a few tens of percent. The sound speed in the upper, rarefied portions of the disk is higher than  $\tilde{c}_{\text{crit}}$ : in a sense only the dense, central layers of the disk are gravitationally stable, and the opaque, upper disk atmosphere convects away the heat dissipated in the center. The temperature is strongly buffered by the two-phase equilibrium: it decreases only a factor  $\sim 1.5$  from midplane to photosphere, even though the pressure drops by several orders of magnitude. The height of the photosphere is  $\sim 10^8$  cm near Earth, so Earth is partially enveloped in a thick silicate haze and appears somewhat like a brown dwarf star.

The main uncertainty in calculating the spreading time of the disk is the photospheric opacity,  $\kappa$ . For a plane-parallel atmosphere in local thermodynamic equilibrium, in a (constant) gravitational field  $g$ , the photospheric pressure  $P_p = g/\kappa = \Omega^2 h_p / \kappa$ . The opacity is probably dominated by silicate grains and depends on the size spectrum of the grains (see, e.g., Fig. 4 of Pollack, McKay, and Christofferson 1985 and the calculations of Alexander, Johnson, and Rypma 1983). If most of the mass is in grains and most of the grains have a diameter comparable to an IR wavelength, then the opacity is maximized and could be as high as  $10^2 \text{ cm}^2 \text{ g}^{-1}$ . More realistically, neither of these requirements is likely to be well satisfied, and we adopt a (very uncertain) nominal value of  $1 \text{ cm}^2 \text{ g}^{-1}$ . Fortunately, our conclusions are not very sensitive to this choice.

One may estimate the photospheric temperature by noting that it depends only logarithmically on the pressure:

$$T_p = T_0 \left[ \ln \left( \frac{P_0}{P_p} \right) \right]^{-1} = 2200 \text{ K} \left\{ 1 + 0.036 \ln \left( \frac{\kappa}{1 \text{ cm}^2 \text{ g}^{-1}} \right) + 0.036 \ln \left[ \frac{h_p}{r_\oplus} \left( \frac{r}{r_\oplus} \right)^3 \right] \right\}^{-1}. \quad (41)$$

The vertical profiles in Figures 7 and 8 show that the height of the photosphere is  $h_p \sim 0.2r_\oplus$  at  $r = 2r_\oplus$ .

The height of the photosphere  $h_p$  may also be estimated following the discussion in § III. The solution  $h_p(r)$  of equation (18) is plotted in Figure 9; it is larger by a factor of 2 than the photospheric heights calculated from the vertical integrations (Table 1). In this approximation,  $h_p$  is independent of  $\sigma$ . We substitute  $z = h_p(r)$  into equation (41) and plot the resulting photospheric temperature  $T_p(r)$  in Figure 4. Note that  $T_p$  is consistently smaller than the mid-plane temperature calculated above.

#### d) Spreading of a Protolunar Disk

The viscous evolution time of the disk is very short. Assuming that the disk radiates as a blackbody, the photospheric temperature on each side of a Keplerian disk is

$$\sigma_{\text{SB}} T_p^4 = \frac{3}{8} v \sigma \Omega^2. \quad (42)$$

The spreading time is therefore:

$$t_{\text{visc}} \sim \frac{r_d^2}{v} \sim \frac{M_d \Omega^2 (r_d)}{\pi \sigma_{\text{SB}} T_p^4} = 100 \left( \frac{r_d}{3r_\oplus} \right)^{-3} \left( \frac{T_p}{2000 \text{ K}} \right)^{-4} \left( \frac{M_d}{2M_\oplus} \right) \text{ yr}. \quad (43)$$

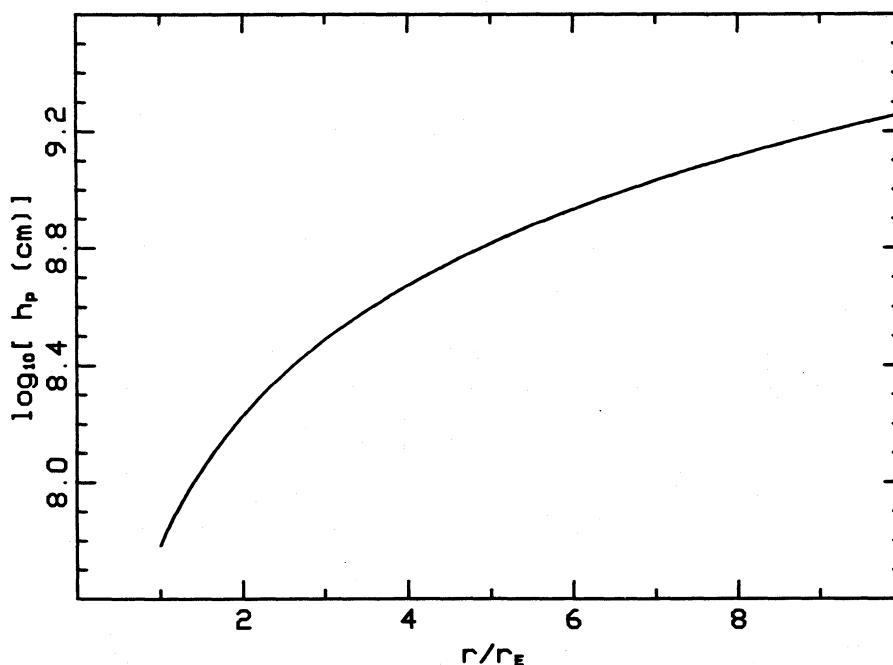


FIG. 9.—Radial dependence of the photospheric height of a hot silicate disk in Earth's orbit

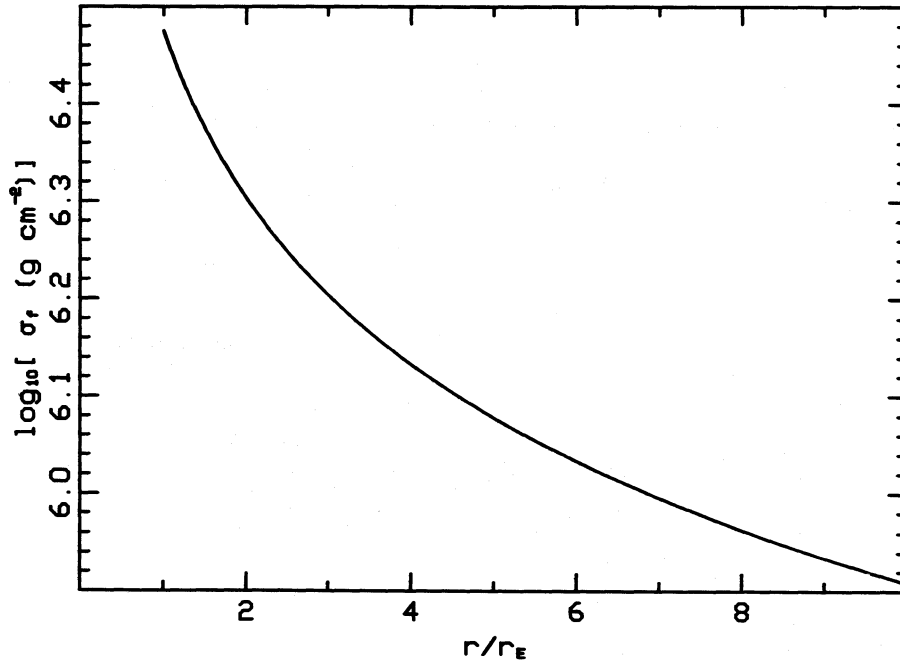


FIG. 10.—Radial dependence of the critical surface density  $\sigma_c$  of a hot silicate disk in Earth orbit, below which gravitational instabilities grow in less than an orbital period (eq. [45]).

Here  $r_d$  is the outer radius of the disk and  $M_d$  is its mass. This time scale is far less than that required to assemble Earth ( $\sim 10^8$  yr).

Equivalently, one may calculate the  $\alpha$  parameter, defined as the ratio of the viscous stress due to differential rotation, to the internal pressure of the disk (Shakura and Sunyaev 1973):

$$\alpha \equiv -\frac{\nu \rho(0) r (d\Omega/dr)}{P(0)} = \frac{4\sigma_{\text{SB}} T_p^4 \mu}{3\sigma \Omega k T(0)} = 6 \times 10^{-5} \left(\frac{r}{3r_{\oplus}}\right)^{3/2} \sigma_7^{-1} \left(\frac{T_p}{2000 \text{ K}}\right)^4 \left[\frac{T(0)}{3000 \text{ K}}\right]^{-1}, \quad (44)$$

for a Keplerian disk. The photospheric temperature would have to exceed  $10^4$  K for  $\alpha$  to reach unity. (As the next calculation shows,  $\sigma$  cannot be less than  $\sim 10^6$  g cm $^{-2}$  if the disk is to remain in this marginally unstable state.)

One may now estimate the critical surface density  $\sigma_c$  at which instabilities grow on an orbital time scale and the disk fragments into individual objects of mass  $\Delta M_f = \sigma_c l_{\text{crit}}^2(\sigma_c)$ . When  $\eta - 1 \approx 1$ , equation (24) gives  $\nu \approx \pi^4 G^2 \sigma^2 \Omega^{-3}$ . Solving for the critical surface density in equation (42), one finds

$$\sigma_c = 1.5 \times 10^6 \left(\frac{r}{r_{\oplus}}\right)^{-1/2} \left\{ 1 + 0.036 \ln \left( \frac{\kappa}{1 \text{ cm}^2 \text{ g}^{-1}} \right) + 0.036 \ln \left[ \frac{z}{r_{\oplus}} \left( \frac{r}{r_{\oplus}} \right)^{-3} \right] \right\}^{-4/3} \text{ g cm}^{-3}, \quad (45)$$

and

$$\Delta M_f = \frac{\pi^4 G^2 \sigma_c^3}{\Omega^4} = 6.2 \times 10^{17} \left(\frac{r}{r_{\oplus}}\right)^{9/2} \left\{ 1 + 0.036 \ln \left( \frac{\kappa}{1 \text{ cm}^2 \text{ g}^{-1}} \right) + 0.036 \ln \left[ \frac{z}{r_{\oplus}} \left( \frac{r}{r_{\oplus}} \right)^{-3} \right] \right\}^{-4} \text{ g}. \quad (46)$$

Notice that at a given radius,  $\sigma_c$  and  $M_f$  depend only on the photospheric opacity. We substitute  $z = h_p(r)$  from Figure 9 and plot the resulting  $\sigma_c(r)$  and  $\Delta M_f(r)$  in Figures 10 and 11.

Upon condensation to a density  $\sim 3$  g cm $^{-3}$ , a lump of  $10^{18}$  g has a radius of 4 km. At a distance of  $10r_{\oplus}$ , the disk fragments into lumps of mass  $\sim 10^{22}$  g and radius  $\sim 100$  km. The  $10^{18}$  g lumps have a surface cooling time of  $\sim 0.1$  yr at a temperature of 2000 K, while the  $10^{22}$  g lumps have a surface cooling time of a few years at the same temperature. If these lumps have composition similar to the bulk Moon, then they may form a "rind" of buoyant crust and be inhibited from internal cooling by the slow diffusion of heat through this crust (closely analogous to the classic Stefan problem; see Turcotte and Schubert 1982, p. 168). However, collisions among lumps rupture these crusts and greatly enhance cooling. At the periphery of the disk, where the thermal bath of silicate vapor has diminished, the source material may thus partially solidify before aggregation into a proto-Moon.

In general, the outer parts of the disk fragment first. We will consider only the simplest case where  $\sigma$  decreases with radius, except near the inner edge, where the disk thins in a turbulent boundary layer (see the discussion below). As the disk spreads, the boundary strip between continuous and fragmented disk moves. Let  $r_f$  be the mean radius of this strip; then

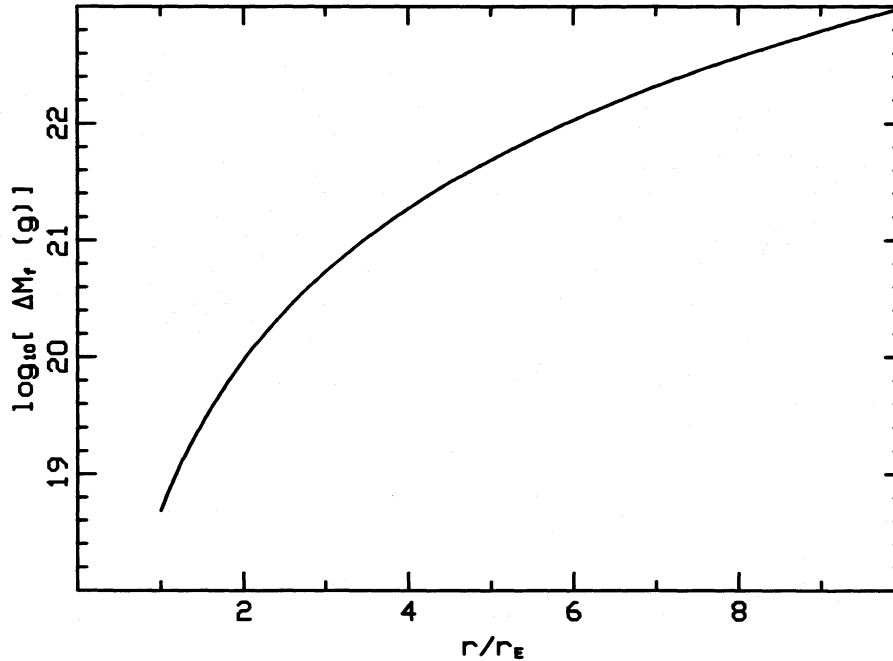


FIG. 11.—Radial dependence of the mass  $M_f$  of the clumps formed, when a hot silicate disk in Earth orbit fragments (eq. [46])

Let  $r_f$  be the mean radius of this strip; then

$$\frac{dr_f}{dt} = -\frac{\partial \sigma}{\partial t}(r_f) \left( \frac{d\sigma_f}{dr} \right)^{-1}. \quad (47)$$

Since  $d\sigma_f/dr < 0$ , one requires the local surface mass density to *grow* with time if  $r_f$  is to increase; this may be possible to arrange during the early stages of spreading, but not asymptotically.

In the simplest case, where fragmentation first appears at a radius  $r_{f,0}$ , and  $r_f$  moves steadily inward until it reaches the Roche radius  $r_{\text{Roche}}$ , one may obtain a maximum value for  $r_{f,0}$  as follows. The rate at which disk material fragments is

$$\frac{dM_f}{dt} = -2\pi r_f \sigma_f(r_f) \left[ \frac{dr_f}{dt} - u_r(r_f) \right], \quad (48)$$

where the velocity  $u_r(r_f)$  of the fragmenting disk material is directed outward. Without calculating  $u_r(r_f)$  from an explicit disk model, we can conclude

$$M_f \geq \int_{r_{\text{Roche}}}^{r_{f,0}} 2\pi r \sigma_f(r) dr = \frac{4\pi}{3} \sigma_0 r_{\oplus}^2 \left[ \left( \frac{r_{f,0}}{r_{\oplus}} \right)^{3/2} - \left( \frac{r_{\text{Roche}}}{r_{\oplus}} \right)^{3/2} \right]. \quad (49)$$

Here  $\sigma_0 \approx 1.5 \times 10^6 \text{ g cm}^{-2}$  (eq. [45]). Requiring  $M_f = 7.4 \times 10^{25} \text{ g}$  (one Moon mass), we obtain  $r_{f,0} \leq 10r_{\oplus}$ .

#### e) Rainout of the Condensed Phase

At this point, we pause to calculate the velocities associated with rotation, sound, convection, and turbulence in a protolunar disk. The orbital velocity around the Earth is

$$v_{\text{orb}} = 7.9 \times 10^5 \left( \frac{r}{r_{\oplus}} \right)^{-1/2} \text{ cm s}^{-1}. \quad (50)$$

In the mixing-length approximation, and neglecting condensation, the vertical convective velocity in the upper regions of the disk is

$$v_{\text{con}} = \left( \frac{l g_z F}{2\rho C_p T} \right)^{1/3}. \quad (51)$$

Here  $l$  is the mixing length, which we take to be  $h_p/10$ ,  $g_z \approx \Omega^2 z$  is the vertical component of gravity,  $F = \sigma_{\text{SB}} T_p^4$  is the heat flux above the central layer where mass and viscous dissipation is concentrated, and  $C_p = (1-x)C_{p1} + xC_{p2}$  is the specific heat of the two-phase fluid. In Figures 7 and 8, we plot  $v_{\text{con}}$  together with the local sound speed. Notice that in the uppermost part of the disk, the convective velocity exceeds the sound speed: we have probably overestimated the mixing length in this region, but in any case we expect this part of the disk to be shock-heated.



The characteristic turbulent velocity due to the disk instabilities is  $v_{\text{turb}} \sim \Omega l_{\text{crit}}(\eta - 1)$ . Combining equations (24) and (42), we may estimate  $\eta - 1$ :

$$(\eta - 1)^2 = \frac{4}{9\pi^4} \frac{\Omega}{G^2 \sigma^3} \sigma_{\text{SB}} T_p^4. \quad (52)$$

At  $r = 2r_{\oplus}$  in a disk with  $\sigma = 10^7 \text{ g cm}^{-2}$ , we obtain  $\eta - 1 = 0.02$  and  $v_{\text{turb}} = 300 \text{ cm s}^{-1}$ . This velocity applies to the dense, central regions of the disk where most of the dissipation takes place.

Given the above estimates, we can address the efficiency of rainout in a convecting disk. The total force on a spherical drop of liquid, with vertical velocity  $v_{1z}$ , suspended in vapor with vertical velocity  $v_{2z}$  is

$$\frac{4\pi}{3} \rho_1 a^3 \dot{v}_{1z} = -\frac{4\pi}{3} \rho_1 a^3 \Omega^2 z + \gamma a^2 \rho_2 (v_{1z} - v_{2z})^2. \quad (53)$$

Here  $\gamma$  is a dimensionless geometrical factor representing large Reynolds number drag. The effects of viscosity may be neglected. The steady state vertical velocity of the drop is

$$|v_{1z} - v_{2z}| \approx \sqrt{\frac{4\pi}{3\gamma} a \Omega^2 z \frac{\rho_1}{\rho_2}}, \quad (54)$$

In order for upward-moving fluid to rain out, we require  $(4\pi/3\gamma)a\Omega^2 z > (\rho_2/\rho_1)v_{2z}^2$ . For example, at  $z = h_{\text{mass}} = 0.07h_p$  in the vertical solution corresponding to  $\sigma = 10^7 \text{ g cm}^{-2}$  and  $r = 2r_{\oplus}$  (Fig. 7),  $v_{2z} = v_{\text{con}} = 100 \text{ cm s}^{-1}$ , and the liquid component rains out only if the "drop" size exceeds  $a \sim 10 \text{ cm}$ .

We may place an upper limit on  $a$  by taking into account the stabilization of a moving drop by its surface tension  $\Sigma$ . The drop is held together if the ram pressure  $\rho_2(v_{1z} - v_{2z})^2 < 2\Sigma/a$ . Substituting  $a = (2\Sigma/\rho_2)(v_{1z} - v_{2z})^{-2}$  in equation (54), we obtain for the maximum drift velocity

$$v_{\text{drift}} = |v_{1z} - v_{2z}| = \left( \frac{4\pi}{3\gamma} \frac{\Omega^2 z \Sigma \rho_1}{\rho_2^2} \right)^{1/4}. \quad (55)$$

We plot this velocity together with the convective velocity in Figures 7 and 8. For all the values of  $\sigma$  and  $r$  tested, the maximum drift velocity is larger than  $v_{\text{con}}$  by a factor 1.5–3, so that rainout can take place. However, a modest decrease in the size of the liquid drops would prevent this. If the disk forms via the settling of a cloud of ejecta, then the estimate (256) for the initial drop size is small enough that the liquid component initially does not rain out.

We still must consider the possibility, raised in the previous section, that the liquid component may settle to the mid-plane in a marginally unstable disk. If this does happen, then the vapor mass fraction and the sound speed fall dramatically at the center of the disk, and the disk breaks up. We saw that the orbital binding energy liberated by the spreading of the disk is sufficient to restore the vertical potential energy liberated by settling if (eq. [19])

$$f \equiv \frac{t_{\text{sed}}}{t_{\text{visc}}} \frac{\mu \Omega^2 r^2}{x k T} > 1. \quad (56)$$

We evaluate  $f = f_{\text{min}}$  in Figure 12 by substituting the minimum possible settling time  $t_{\text{sed}} = h_{\text{mass}}/v_{\text{drift}}$ . For  $v_{\text{drift}}$  we use equation (55), unless this velocity is larger than that of free Keplerian orbits through the disk. In this second case, we use instead  $t_{\text{sed}} = \Omega^{-1}$ . The break in the plotted values of  $f_{\text{min}}$  occurs where these two velocities are equal. For the viscous spreading time, we use  $t_{\text{visc}} = r^2/\nu = (9/8)\sigma r^2 \Omega^2 (\sigma_{\text{SB}} T_p^4)^{-1} = 7.2 \times 10^9 \text{ s } \sigma_7 (r/r_{\oplus})^{-1} (T_p/2000 \text{ K})^{-4}$ . Finally, we choose  $z = h_{\text{mass}} = c_{\text{crit}}/\Omega = \pi G \sigma / \Omega^2$  and  $x = \tilde{x}_{\text{crit}}$  (eq. [34]). When  $\sigma = 10^7 \text{ g cm}^{-2}$ ,  $f_{\text{min}}$  ranges from 0.03 to 0.5, for  $r_{\oplus} < r < 10r_{\oplus}$ . When  $\sigma = 10^6 \text{ g cm}^{-2}$ ,  $f_{\text{min}}$  ranges instead from 0.3 to 1.6.

We conclude that some overall settling of the liquid component can take place. However, this settling does not cause the disk to be so unstable that it breaks up. The disk remains marginally unstable, as we now argue.

It was our estimate (41) of the photospheric temperature which led us to conclude that the protolunar disk was only marginally unstable (eq. [52]). The dissipation in the disk can increase substantially before instabilities in the disk grow on an orbital time and the disk fragments. For example, we see from equation (52) that the dissipation rate is  $\sim 3000$  times larger when  $\eta - 1 \sim 1$ , than when the disk sits in the marginally unstable state, for  $\sigma = 10^7 \text{ g cm}^{-2}$  and  $r = 3r_{\oplus}$ . In comparison, only a factor of 10 increase in the dissipation rate is required to prevent settling.

This leads us to enumerate three possible evolutionary paths for a hot protolunar disk that starts out with a surface density  $\sim 10^7 \text{ g cm}^{-2}$ . First, the condensed phase will not settle out if the drop size of this phase is small enough, and the disk behaves as we have outlined. Second, settling of the liquid component may force the disk into a more dissipative state, so that it spreads in 1–10 yr, instead of 100 yr, but still remains marginally unstable. Third, the disk may undergo brief outbursts, in which the dissipation increases perhaps by a factor of 100, thoroughly mixing the two-phase medium and breaking up the liquid component into smaller fragments. It seems that the evolution of the disk depends on details that are much too complicated to reproduce in any conceivable numerical simulation. Nonetheless, we do reach the firm conclusion that a disk massive enough to form the Moon must have been hot and spread very quickly.

#### f) Realignment of the Disk

There is no reason to expect that after a giant impact, the spin angular momentum of the Earth is aligned with the disk angular momentum. The two will eventually align, since parts of the disk at different distances from the Earth precess at different rates,

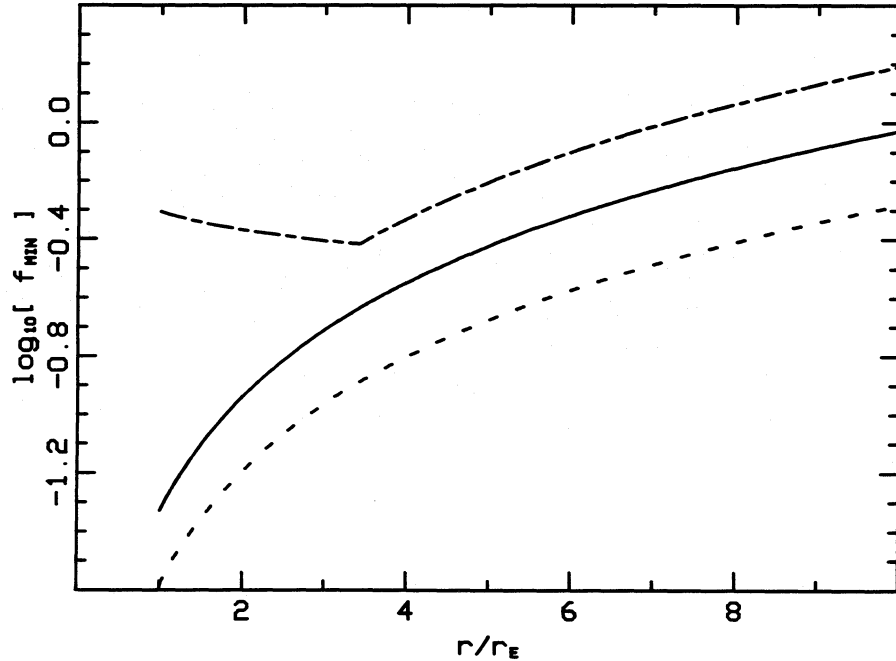


FIG. 12.—Radial dependence of the quantity  $f_{\min}$  (eq. [56]), defined so that when  $f_{\min} > 1$ , the viscous heating of the disk is sufficient to prevent settling of the liquid component to the mid-plane. The three curves correspond to different values of the surface density:  $10^7 \text{ g cm}^{-2}$  (dashed);  $3 \times 10^6 \text{ g cm}^{-2}$  (solid);  $10^6 \text{ g cm}^{-2}$  (double-dashed).

exciting spiral bending waves. The energy flux in these waves through the outer edge of the disk is dissipated, and once the waves are damped, the two angular momenta are aligned. Disk segments precess because Earth's potential has a quadrupole moment, which is produced by Earth's rapid rotation. Indeed, most of the angular momentum of the Earth-Moon system is in Earth's spin [ $I_{\oplus} = 8 \times 10^{44} \text{ g cm}^2 \text{ s}^{-2}$ ,  $r_{\ell}$  (today)  $= 3.8 \times 10^{10} \text{ cm}$ ]:

$$J_{\text{tot}} = I_{\oplus} \Omega_{\oplus} + M_{\ell} (GM_{\oplus} r_{\ell})^{1/2}. \quad (57)$$

Here we have made the approximation that the angular momenta of Earth and Moon are aligned during the tidal evolution of the Moon away from Earth, and neglected the spin angular momentum of the Moon.

Taking  $r_{\ell} = 2r_{\oplus}$ , we find  $\Omega_{\oplus} = 3.7 \times 10^{-4} \text{ s}^{-1} = 2\pi(0.2 \text{ day})^{-1}$ , and  $J_{\oplus}/J_{\ell} = 5.5$ . (Equivalently,  $\Omega_{\oplus} [GM_{\oplus}/r_{\oplus}^3]^{-1/2} = 0.089$ .)

Including a quadrupole term, Earth's potential may be written

$$\Phi = -\frac{GM_{\oplus}}{r} + \Phi_2 = -\frac{GM_{\oplus}}{r} \left[ 1 - J_2 \left( \frac{r_{\oplus}}{r} \right)^2 P_2(\cos \theta) \right]. \quad (58)$$

The value of the gravitational moment  $J_2$  depends partly on the rotation rate and partly on the internal structure of Earth. In general,  $J_2$  is related to the ratio of centrifugal to gravitational accelerations at the equator by the quadrupole response coefficient  $\Lambda_2$ , which is a number of order unity:

$$J_2 = \Lambda_2 \frac{\Omega_{\oplus}^2 r_{\oplus}}{GM_{\oplus}/r_{\oplus}^2}. \quad (59)$$

(See, for example, Hubbard 1984, p. 95.) For uniform density,  $\Lambda_2 = 0.5$ . For the real Earth,  $\Lambda_2 \approx 0.3$ , using the Radau-Darwin approximation (Hubbard 1984).

In the limit  $J_{\oplus} \gg J_{\text{disk}}$ , we can assume that Earth's angular momentum remains fixed during the realignment of the disk. We estimate the energy flux of bending waves near the outer edge of the disk to be

$$\dot{E} = 2\pi r_d \sigma v_g [\Delta v(r_d)]^2, \quad (60)$$

where  $v_g$  is the group velocity of the waves and  $\Delta v$  is the vertical velocity of the disk due to the bending waves. The dispersion relation for axisymmetric bending waves in a Keplerian disk is  $\omega^2 = \Omega^2 + 2\pi G \sigma k$  (see, e.g., Shu 1985). In contrast, the dispersion relation for the bending waves excited in a differentially precessing disk is similar to that for linear ripples in a self-gravitating sheet:  $\omega^2 = 2\pi G \sigma k$ . This is because an axisymmetric bending wave perturbs disk material out of its orbital plane (hence the vertical restoring force proportional to  $\Omega^2$ ), whereas in a differentially precessing disk, the deviation of fluid particles from their orbital plane is very small over one orbital period. Note that the disk self-gravity provides a positive restoring force for both types of bending wave, so that instabilities similar to those discussed above for density waves do not occur.

We must distinguish between two cases, depending on whether  $M_d$  is greater or less than the mass associated with the quadrupole piece of Earth's potential,  $M_Q \equiv \Omega_\oplus^2 r_\oplus^3 / G$ . If a flat disk is initially aligned with respect to Earth's equatorial plane, then differential precession can be viewed as pulling material out of the initial disk plane. The force per unit mass responsible for this is

$$F_Q = -\frac{1}{r} \frac{\partial \Phi_Q}{\partial \theta} \left( \theta = \frac{\pi}{2} - \epsilon \right) \approx -\epsilon r \times \Omega_\oplus^2 \left( \frac{r}{r_\oplus} \right)^{-5}, \quad (61)$$

where the angular distance of disk material from Earth's equatorial plane is initially  $\epsilon = (\pi/2 - \theta) = \epsilon_0 \cos(\phi)$ . The restoring force on the rippled disk due to its self-gravity is

$$F_d = \Delta\epsilon r \times 2\pi G \sigma k, \quad (62)$$

where  $\Delta\epsilon$  is the angular deviation of the material from the initial disk plane. If  $F_d \gg F_Q$  at  $\Delta\epsilon \sim \epsilon$ , then differential precession is inhibited, and only long-wavelength warps in the disk occur; the inner parts of the disk precess at a rate comparable to the outer parts, more slowly than they would in a low-mass disk. At the other extreme, different parts of the disk precess independently, and bending waves are excited, some of which propagate outward and are dissipated in the outer atmosphere of the disk. The boundary between these two types of behavior is

$$2\pi G \sigma k = \Omega_\oplus^2 \left( \frac{r}{r_\oplus} \right)^{-5}, \quad (63)$$

or, equivalently,

$$M_d = \left( \frac{1}{4\pi\beta} \right) \left( \frac{r_d}{r_\oplus} \right)^{-2} M_Q, \quad (64)$$

where  $k$  is a few times  $2\pi/r$ ,  $k = \beta 2\pi/r$ . Since  $M_Q \approx 0.09 M_\oplus$ , the ratio of these two masses is  $M_d/M_Q = 0.25$  for  $M_d = 2M_\oplus$ . The numerical factor in the above equation is  $\sim 0.01$ , however, so the the protolunar disk is massive enough to suppress partially its differential rotation. By setting  $F_d = F_Q$ , we estimate

$$\frac{\Delta\epsilon}{\epsilon} \sim \frac{1}{4\pi^2} \frac{\Omega_\oplus r_\oplus}{G\sigma} \left( \frac{r}{r_\oplus} \right)^{-4} = 3.3 \left( \frac{\sigma}{10^7 \text{ g cm}^{-2}} \right)^{-1} \left( \frac{r}{r_\oplus} \right)^{-4}, \quad (65)$$

since in this case  $k \sim 2\pi/r$ . It is only in the very innermost segment of the disk that the disk's self-gravity does not inhibit differential precession.

The time averaged vertical velocity of disk material is then

$$(\Delta v)^2 \approx \frac{1}{2} \times 2\pi G \sigma k (\Delta\epsilon_0 r)^2 = \frac{1}{8\pi^2} \epsilon_0^2 (\Omega_\oplus r_\oplus)^2 \left( \frac{\Omega_\oplus^2 r_\oplus}{G\sigma} \right) \left( \frac{r}{r_\oplus} \right)^{-7}, \quad (66)$$

where  $\Delta\epsilon_0 = \epsilon_0(\Delta\epsilon/\epsilon)$  is the maximum angular deviation of the disk material. The group velocity at  $r = r_d$  is

$$v_g = \frac{d\omega}{dk} = \sqrt{\frac{\pi G \sigma}{2k}} \sim \frac{1}{2} \sqrt{G \sigma r_d}. \quad (67)$$

From the virial theorem, the energy difference between a disk at inclination  $\epsilon_0$  and one at zero inclination is

$$\Delta E = \frac{1}{2} \int_{r_\oplus}^{r_d} dr \int_0^{2\pi} d\phi \left[ 2\pi r \sigma \times \frac{1}{2} \epsilon(\phi) \frac{\partial^2 \Phi_Q}{\partial^2 \theta} (\pi/2) \right] = \frac{\pi^2}{2} \epsilon_0^2 \Omega_\oplus^2 \sigma r_\oplus^4 \left( 1 - \frac{r_\oplus}{r_d} \right), \quad (68)$$

where the angular distance of a part of the disk above Earth's equatorial plane is  $\epsilon(\phi) \approx \epsilon_0 \cos(\phi)$ . (We have assumed that  $\sigma$  is constant in performing the integrals.) The alignment time of the disk is then

$$t_{\text{align}} \sim \frac{\Delta E}{\dot{E}} = 2\pi^{3/2} \frac{M_d}{M_Q} \left( \frac{M_\oplus}{M_d} \right)^{1/2} \left( \frac{r_d}{r_\oplus} \right)^3 \left( 1 - \frac{r_\oplus}{r_d} \right) t_{\text{orb}}(r_d) \approx 3 \times 10^{-3} \text{ yr} \left( \frac{r_d}{r_\oplus} \right)^{9/2} \left( \frac{M_d}{2M_\oplus} \right)^{1/2} \quad (r_d \gg r_\oplus). \quad (69)$$

Although this calculation assumed a small disk inclination, the alignment time should be similar for moderate angles. We conclude that the disk is torqued into the equatorial plane of Earth in a time short compared to its spreading time.

The instabilities discussed in § III will exist into inner edge of the disk, which presumably is Earth's surface. Since Earth's equatorial rotation velocity is less than the orbital velocity at that radius, there will be a turbulent boundary layer between Earth and disk. Material in the inner regions of the disk will flow inward, at a rate  $\dot{M} \sim M_d/t_{\text{visc}}$ . Since the flow time across the boundary layer is much less than  $t_{\text{visc}}$ , the flow will be steady on the scale of the boundary layer, and conservation of angular momentum gives

$$\frac{d}{dr} (\dot{M} r^2 \Omega) = \frac{d}{dr} \left( -2\pi \sigma v r^3 \frac{d\Omega}{dr} \right). \quad (70)$$

(See, e.g., Pringle 1981.) That is, there is a balance between the inward convective flux of angular momentum and the outward viscous flux. Using the fact that  $\Omega$  has a maximum near  $r_\oplus$ , we can write

$$\dot{M}r^2\Omega \approx -2\pi\nu\sigma r^3 \frac{d\Omega}{dr} + \dot{M}r_\oplus^2 \Omega(r_\oplus). \quad (71)$$

Here we distinguish  $\Omega(r_\oplus)$ , the orbital angular velocity at  $r_\oplus$ , from  $\Omega_\oplus$ , Earth's angular velocity of rotation. Approximately  $d\Omega/dr$  at Earth's surface as  $\Omega(r_\oplus)/\Delta r$  (where  $\Delta r$  is the radial thickness of the boundary layer), we find

$$v_{bl} \sigma_{bl} \approx \frac{1}{2\pi} \dot{M} \frac{\Delta r}{r_\oplus}. \quad (72)$$

Here  $v_{bl}$  and  $\sigma_{bl}$  are the viscosity and surface density in the boundary layer. The Reynolds number of the boundary layer flow is  $\Delta v \Delta r / \nu_{bl} \sim 10^3$ , where  $\Delta v = [\Omega(r_\oplus) - \Omega_\oplus]r_\oplus \approx \Omega(r_\oplus)r_\oplus$ , so that

$$\sigma_{bl} \approx 10^3 \frac{\dot{M}}{2\pi r_\oplus^2 \Omega(r_\oplus)}. \quad (73)$$

Using  $\dot{M} = M_\ell / t_{\text{visc}}$ , with  $t_{\text{visc}} = 300$  yr (see eq. [43]), we find  $\sigma_{bl} = 2500 \text{ g cm}^{-2}$  and  $h_{\text{mass, bl}} > 10$  m. (The turbulent velocities are  $v_{\text{turb}} \sim v_{bl}/\Delta r \sim 10^{-3}\Omega(r_\oplus)r_\oplus$ —that is, subsonic.) In a circular turbulent boundary layer, the thickness  $\Delta r \sim r(v_{\text{turb}}/\Delta v)$ , which means that in our case  $\Delta r \sim 10^{-3}r_\oplus \sim 10 \text{ km} \gg h_{\text{mass, bl}}$ . The angular momentum of the disk material accreted onto Earth is too small to affect significantly Earth's spin.

#### g) Mass Loss from the Disk

Because the protolunar disk is in a very hot, turbulent state, it is interesting to consider whether a wind may be driven off it. A reasonable value for the outer radius of the disk when it begins to fragment is  $10r_\oplus$ ; at this radius and at a temperature of 2000 K, a molecule of atomic weight 30 has a significant fraction (0.13) of the escape energy. Mass loss through the disk “exosphere” is not interesting on a (very short) viscous spreading time, however. Rather, it is possible that acoustic waves propagating into the hot disk atmosphere may drive a wind. Since the turbulence velocities in the upper, rarefied layers of the disk are close to the sound speed, a rough estimate of the energy input into the disk atmosphere is

$$\dot{E} = 2\pi r_d^2 \rho_p c_p^3. \quad (74)$$

The corresponding mass loss is (if this energy is channeled efficiently into a wind)

$$\dot{M}_{\text{visc}} = \frac{2\pi\rho_p c_p^3}{\Omega^2(r_d)} \times \frac{M_d \Omega^2(r_d)}{\sigma_{\text{SB}} T_p^4} = 7 \times 10^{-4} M_d \left( \frac{\rho_p}{10^{-7} \text{ g cm}^{-3}} \right) \left( \frac{c_p}{10^4 \text{ cm s}^{-1}} \right)^3 \left( \frac{T_p}{2000 \text{ K}} \right)^{-4}. \quad (75)$$

This is likely to be a small fraction of the disk mass, if the escaping gas consists of SiO or O<sub>2</sub> molecules. It could be more substantial if it is hydrogen, derived from the dissociation of water. It seems probable that the disk is capable of efficient devolatilization, at least to the extent of water loss. We doubt that there is any possibility of efficient fractionation of the type hypothesized by Cameron (1983) for xenon isotopes, because the loss is hydrodynamic. We should emphasize that the mass loss is greater than the estimate (75) if the dissipation rate in the disk is higher (see § IVe).

#### h) Fragmentation of the Disk

Finally, we raise a question that is central to any disk-formation model for the origin of the Moon: What fraction of the disk mass enters the Moon, or correspondingly what fraction accretes onto Earth? The time-dependent numerical simulation required to answer this question is beyond the scope of this paper. On purely energetic grounds, it is possible for most of the disk mass to spread well beyond the Roche radius: the binding energy per unit mass released by inflowing material near Earth's surface is greater than that absorbed by outflowing material at the Roche radius.

In the disk model which we have investigated in detail, the temperature of the photosphere is only very weakly dependent on the local surface density. We now show that this model cannot provide a complete description of the outer parts of the disk.

The slow variation of the photospheric temperature implies some unusual properties of the disk viscosity. Since  $\nu\sigma \sim r^3 T_p^4$  (from eq. [42]), the viscosity increases with radius, probably very quickly. Indeed, the local spreading time is proportional to  $r^2/\nu$  and decreases with radius so long as  $\sigma/r$  decreases with radius. This indicates that the outer parts of the disk, which presumably contain most of the mass, evolve more rapidly than the inner parts.

An inconsistency in the model may be found by considering the evolution of the disk. The equations of mass and angular momentum conservation in the disk may be convolved to obtain the following differential equation (Pringle 1981):

$$\frac{\partial\sigma}{\partial t} = \frac{3}{r} \frac{\partial}{\partial r} \left[ r^{1/2} \frac{\partial}{\partial r} (\nu\sigma r^{1/2}) \right]. \quad (76)$$

If  $T_p$  is constant, then

$$\frac{\partial\sigma}{\partial t} = \frac{63}{2} \frac{\nu\sigma}{r^2}, \quad (77)$$



and the density increases locally. But clearly this equation cannot hold everywhere in the disk—if it did, then mass would not be conserved. Further insight is provided by calculating the radial velocity of disk material:  $r\sigma u_r = -3r^{1/2}(\partial/\partial r)(v\sigma r^{1/2})$ . If  $T_p$  is constant, then

$$r\sigma u_r = -\frac{21}{2}v\sigma. \quad (78)$$

Marginally unstable disk material always moves *inward* toward Earth. But conservation of angular momentum requires that the outer parts of the disk move away from Earth. In the outer parts of the disk, the scaling of  $v\sigma$  with radius must be substantially different from that obtained from the assumption that the photosphere regulates the dissipation in the disk.

Thus, it is difficult to see how the disk could remain marginally unstable in its outer parts. If the photosphere were in two-phase equilibrium, then in order for  $T_p$  to decrease as  $r^{-3/4}$  or faster, the photospheric opacity  $\kappa$  would have to increase *exponentially fast* with radius; see equation (41). Alternately, if the photosphere were not in two-phase equilibrium, then the buffer preventing the disk from becoming more unstable would be removed, and one would expect the viscosity to increase rather than decrease.

A resolution of this apparent paradox is provided by a careful treatment of the energy balance within the disk. Recall that the losses from the photosphere are a combination of viscous dissipation and secular cooling (see eq. [23]). Here we concentrate on the latent heat released when the vapor phase condenses. Once the disk has settled into the marginally unstable state, the vertical mass average

$$\tilde{x}^2 = \frac{1}{\sigma} \int_{-h}^h x^2 \rho(z) dz$$

is essentially a function of  $\sigma$  and  $r$  (see eq. [34]). Indeed,  $\tilde{x}$  in a marginally unstable disk is proportional to  $\sigma$ , so that latent heat is released as the surface density decreases.

The surface density in the vapor phase is  $\langle x \rangle \sigma$ , where

$$\langle x \rangle = \frac{1}{\sigma} \int_{-h}^h x \rho(z) dz. \quad (79)$$

Now, the vertical disk profiles of Fig. 7 show that  $x$  is almost constant in the central layers of the disk which contain most of the mass. We then have  $\langle x \rangle \approx \tilde{x}$ , and the equation describing conservation of internal energy in the disk is approximately

$$2\sigma_{\text{SB}} T_p^4 = \frac{9}{4} v\sigma\Omega^2 - \left[ \frac{\partial}{\partial t} (l\sigma\tilde{x}) + \frac{1}{r} \frac{\partial}{\partial r} (u_r r l\sigma\tilde{x}) \right]. \quad (80)$$

Here  $l$  is the latent heat of condensation per unit mass. Making use of equation (34) and the equation of continuity, we may write the expression in brackets in equation (80) as

$$\sigma \left( \frac{\partial \sigma}{\partial t} + u_r \frac{\partial \sigma}{\partial r} + \frac{3}{2} \frac{\sigma u_r}{r} \right) \frac{\pi G}{\Omega} \sqrt{C_{p1} T(0)}. \quad (81)$$

It is convenient to reexpress the derivatives of  $\sigma$  in terms of derivatives of  $v\sigma$ , after which equation (80) takes the form

$$2\sigma_{\text{SB}} T_p^4 = \frac{9}{4} v\sigma\Omega^2 - \sigma \left\{ \frac{\partial^2 (v\sigma)}{\partial r^2} - \frac{3}{4} \frac{v\sigma}{r^2} - \frac{1}{\sigma} \frac{\partial \sigma}{\partial r} \left[ \frac{\partial (v\sigma)}{\partial r} + \frac{1}{2} \frac{v\sigma}{r} \right] \right\} \frac{3\pi G}{\Omega} \sqrt{C_{p1} T(0)}. \quad (82)$$

Note that the latent heat has dropped out of equation (82). If all quantities in equation (82) scale as powers of  $r$ , then the heat released by condensation is a small fraction of the viscous dissipation:

$$\frac{\dot{E}_{\text{conden}}}{\dot{E}_{\text{visc}}} \sim \frac{\pi G \sigma}{\Omega^3 r^2} \sqrt{C_{p1} T(0)} = 7 \times 10^{-3} \left( \frac{r}{r_{\oplus}} \right)^{5/2} \left( \frac{r_d}{r_{\oplus}} \right)^{-2} \left( \frac{M_d}{2M_{\oplus}} \right) \left[ \frac{T(0)}{3000 \text{ K}} \right]^{1/2}. \quad (83)$$

(Here we have taken  $M_d = \pi r_d^2 \sigma$ .)

At first sight, it would appear that the release of latent heat is never an important contribution to the energy budget. However,  $\dot{E}_{\text{conden}}$  is enhanced with respect to  $\dot{E}_{\text{visc}}$  if  $\sigma$  and  $v\sigma$  decrease rapidly, over a length scale  $l \ll r$ .

We wish to consider a particular model, in which the radial flow direction changes sign in a thin boundary layer at the outer edge of the disk. Since  $r\sigma u_r = -3r^{1/2}(\partial/\partial r)(v\sigma r^{1/2})$ , the viscous dissipation must decrease from the inner to the outer edge of the boundary layer. We assume that the condensation of the vapor phase compensates for the decrease in viscous dissipation; this is why the boundary layer is thin.

Since the disk material interior to the boundary layer is flowing inward, the inner radius of the boundary layer  $r_b$  must also move inward. Indeed,  $r_b$  moves slightly faster than the material interior to it. In a frame comoving with  $r_b$ , there is an *outward* flow of disk material across the boundary layer. On the time scale over which material crosses the boundary layer, this flow is steady. That is, the disk loses mass through the boundary layer at the rate

$$\dot{M}_d = 2\pi r_b \sigma(r_b) [\dot{r}_b - u_r(r_b)] = 2\pi r \sigma(r) [\dot{r}_b - u_r(r)] < 0, \quad (84)$$

where  $u_r(r_b)$  and  $\sigma(r_b)$  are the radial velocity and surface density just interior to  $r_b$ . Combining the equation of continuity with

equation (84), we see that

$$\frac{\partial \sigma}{\partial t} \approx -\frac{\partial \sigma}{\partial r} \dot{r}_b \quad (85)$$

in the boundary layer.

In the boundary layer, expression (81) simplifies to

$$\frac{\partial \sigma}{\partial t} (\dot{r}_b - u_r) \frac{\sigma}{\dot{r}_b} \times \frac{\pi G}{\Omega} \sqrt{C_{p1} T(0)} = 3 \frac{\partial^2 (v\sigma)}{\partial r^2} \frac{|\dot{M}_d|}{2\pi r_b |\dot{r}_b|} \times \frac{\pi G}{\Omega} \sqrt{C_{p1} T(0)}. \quad (86)$$

In the equality, we have made use of equations (84) and (85). Defining

$$2\sigma_{SB} T_p^4 - \frac{9}{4} v\sigma \Omega^2 = \frac{9}{4} \Delta(v\sigma) \Omega^2, \quad (87)$$

the energy conservation equation (80) becomes

$$\frac{\partial^2 \Delta(v\sigma)}{\partial r^2} = \frac{\Delta(v\sigma)}{l_b^2}, \quad (88)$$

in the boundary layer. Here the length scale  $l_b$  may be written

$$\left(\frac{l_b}{r_b}\right)^2 = \left(\frac{4\kappa}{3}\right) \frac{\pi G \sigma(r_b)}{\Omega^3(r_b) r_b^2} \sqrt{C_{p1} T(0)} \approx \frac{\dot{E}_{\text{conden}}}{\dot{E}_{\text{visc}}} \bigg|_{r_b}, \quad (89)$$

since  $|\dot{M}_d| = \kappa 2\pi r_b |\dot{r}_b| \sigma(r_b)$ , to within a factor  $\kappa$  of order unity. Referring to equation (82), we see that the boundary layer is not very thin, namely  $l_b/r_b \sim 0.2$ .

At  $r = r_b$ , we require that  $\Delta(v\sigma) = 0$  and that  $\partial(v\sigma)/\partial r$  be continuous, so the solution to equation (89) is

$$\Delta(v\sigma) = \frac{l_b}{2} \left[ \frac{\partial(v\sigma)}{\partial r} \right]_{r_b} [e^{(r-r_b)/l_b} - e^{-(r-r_b)/l_b}] \quad (r \geq r_b). \quad (90)$$

The surface density profile in the boundary layer is similarly determined, by noting that  $|\dot{r}_b| \partial \sigma / \partial r \approx 3 \partial^2 (v\sigma) / \partial r^2$ :

$$\sigma = \sigma(r_d) - \frac{3}{2|\dot{r}_b|} \left[ \frac{\partial(v\sigma)}{\partial r} \right]_{r_b} [e^{(r-r_b)/l_b} + e^{-(r-r_b)/l_b} - 2] \quad (r \geq r_b). \quad (91)$$

Note that these profiles match smoothly onto the disk interior to  $r_b$ .

The boundary layer solution given by equations (90), (91) is compatible with our understanding of the vertical structure of the disk. Since  $\tilde{x}$  is proportional to  $\sigma$ , it follows that  $\sigma$  must decrease if latent heat is to be released. To satisfy conservation of mass, the radial velocity (as measured in the frame comoving with  $r_b$ ) increases rapidly outward in the boundary layer.

We note that  $\sigma$  decreases much faster than  $v\sigma$  in the boundary layer. That is,  $\sigma$  drops to a value much smaller than  $\sigma(r_b)$  in a distance  $r - r_b \sim l_b$ ; whereas at a similar distance,  $v\sigma$  has decreased only by a fractional amount  $\sim l_b/r_b$ . It is meaningless to extrapolate equations (90) and (91) further than this, since the disk breaks up when  $\sigma$  is comparable to the critical surface density (eq. [45]) at which gravitational instabilities grow on an orbital time scale. Note that our expression (45) for the critical surface density is not altered in the boundary layer, because  $v\sigma$  decreases only slightly.

We conclude that the outer parts of the disk must be continually *fragmented*, provided that they lie beyond the Roche radius. Conservation of angular momentum requires that the marginally unstable inner disk be surrounded by an annulus in which the surface density decreases rapidly outward. Material in this annulus expands outwards and fragments. Just as icebergs break off a contracting ice sheet, so one expects the disk to shed fragments of mass  $\sim \Delta M_f$  given by equation (46). As the boundary between continuous and fragmented disk moves inward,  $\Delta M_f$  decreases.

The main conclusion to be drawn from this discussion is that the marginally unstable disk does not spread outward. Rather, it contracts essentially from the moment it is formed, losing angular momentum to small fragments at its outer edge.<sup>5</sup>

If the disk is initially formed *completely* within the Roche radius, then this picture is modified. The fragments formed at the outer edge of the marginally unstable disk are disrupted, and blend into a semicontinuous outer disk. This disk will spread outward, since it is not in the marginally unstable state. (Here the source of viscosity is similar to that postulated by Ward and Cameron 1978.) Fragments from the outer disk may collect in a proto-Moon just outside the Roche radius. Note that, in this case, most of the disk mass is accreted back onto Earth.

We have not found any sensible way to calculate the amount of disk material incorporated into a satellite. One possibility is that all disk material which fragments outside the Roche radius is incorporated into a satellite. Even if this is the case, the satellite mass depends on the outer radius at which fragmentation begins. The radial velocity of material is directed outward at the outer radius of

<sup>5</sup> Angular momentum may be extracted from the inner disk via collisions between the fragments. If the number of fragments is sufficiently small, then angular momentum may be extracted via tidal interactions. This is analogous to the situation studied by Goldreich and Tremaine (1980), in which a shepherd satellite maintains a sharp outer boundary to a cold disk (such as Saturn's rings) by extracting angular momentum from the disk.



the boundary layer, where the disk fragments; so the lower bound (49) on the mass of the fragments should apply.

Since the marginally unstable disk material has a strong tendency to flow inward, the danger clearly exists that only a small fraction of the disk fails to be accreted back onto the proto-Earth. If the ejecta from a giant impact are processed through a disk before being incorporated into the Moon, then a greater total mass of ejecta may be required than was produced in the simulations of Benz, Slattey, and Cameron (1986, 1987).

## V. CONCLUSION

We have shown that a two-phase disk may be gravitationally unstable at much higher temperatures or lower surface densities than a single-phase disk. Since photospheric losses are limited by the two-phase equilibrium condition, a two-phase disk may enter a metastable, marginally unstable state. The disk fragments only when the local growth time of instabilities required to maintain the photospherically regulated dissipation becomes comparable to the orbital time. In our admittedly speculative application to the impact hypothesis of lunar origin, this leads to the conclusion that the prelunar disk can spread and cool on a time scale of  $\sim 10^2$  yr (although under some circumstances the spreading time might be a factor of 10 shorter than this). The whole disk is hot and convecting, and the mass is concentrated toward the midplane, sandwiched by a rarefied, optically thick silicate atmosphere. The Moon is thereby formed molten, or partially molten. The most significant result of this work is not that it provides partial support for the impact hypothesis (the validity of which rests more on the outcome of the calculations of Benz *et al.* and others in progress), but that it provides a semiquantitative description of the fate of a disk of liquid and vapor in Earth orbit. This description places constraints on the nature of lunar structure and early evolution which are addressed more fully elsewhere (Hartmann, Phillips, and Taylor 1986; Stevenson 1987). Here we briefly enumerate the issues and problems posed by our analysis:

1. Despite the trauma of the impact event and the extensive (presumably incongruent) vaporization of Moon-forming material, the proto-Earth–proto-Moon system is almost closed: very little material leaves, except hydrogen and possibly some other volatiles. Major element fractionation is not likely, and the Moon should be similar to the bulk composition of the material placed in Earth orbit following the impact. Dehydration should be almost complete, however.

2. Although the model demonstrates the spreading of the disk beyond the Roche limit and the accumulation of protomoons, it offers no firm prediction concerning the number and disposition of Moons. We cannot assess at present whether a *single* Moon is guaranteed from this process, neither do we necessarily expect that there was only one disk-forming event. However, the lifetime of the disk is far less than the time between impacts capable of creating (or affecting) a disk, so the events can be treated discretely.

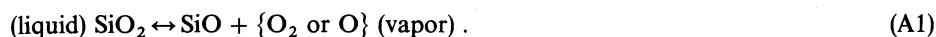
3. The model predicts a hot start for the Moon. This is difficult to achieve in most other theories of lunar origin. We cannot predict with confidence the extent of melting in the finally accumulated body, for the reasons evident from (2) above. Nevertheless, magma oceans on both Earth and Moon seem highly probable.

4. The model predicts that the initial Moon orbital plane coincided with Earth's equatorial plane. This might seem to be inconsistent with backward extrapolations of lunar tidal history (e.g., Goldreich 1966). However, these calculations require a number of assumptions about the frequency dependence, amplitude dependence, and spatial variation of tidal dissipation in the Earth, the validity of which cannot currently be assessed. Of course, it is also possible that the Earth suffered another substantial collision, changing its obliquity, after the Moon-forming event.

We acknowledge conversations with P. Goldreich and W. Ward. This work was initiated while C. T. held a Caltech Summer Undergraduate Research Fellowship; we are grateful to the SURF program for its support. D. J. S. acknowledges support from NASA grant NAGW 185.

## APPENDIX

In this Appendix, we detail our calculation of the sound speed in the following two-phase system:



Recall from § II that the two-phase sound speed is

$$c = \frac{(\rho_2/\rho_1)(1-x) + x}{\sqrt{(C_{p1}T/l^2)(1-x) + [(\mu_2/kT) - (2/l) + (C_{p2}T/l^2)]x}} , \quad (\text{A2})$$

Here  $x$  is the mass fraction in the vapor phase (phase 2). The mean molecular weight in phase 2 is  $\mu_2 = 30$  a.m.u. for a  $\text{SiO} + \frac{1}{2}\text{O}_2$  mixture. To obtain the specific heat in the liquid phase (phase 1) we extrapolate the empirical value for quartz (measured up to  $\sim 1800$  K):

$$C_{p1}(T) = 9.82 \times 10^6 + 1.67 \times 10^3 T \text{ ergs g}^{-1} \text{ K}^{-1} , \quad (\text{A3})$$

(Robie, Hemingway, and Fisher 1978).  $C_{p2} = (3/2)k/\mu$ . We may obtain the heat of vaporization,  $l$ , and the constants  $P_0$  and  $T_0 = l\mu_2/k$  in the phase equilibrium condition  $P(T) = P_0 e^{-T_0/T}$ , from either the calculations of Krieger (1967), or the experiment of Ruff and Schmidt (1921). The corresponding values, for  $\mu_2 = 30$ , are  $l = 1.7 \times 10^{11}$  ergs  $\text{g}^{-1}$ ,  $P_0 = 3.2 \times 10^{14}$  dyne  $\text{cm}^{-2}$ , and  $T_0 = 6.0 \times 10^4$  K (Krieger); and  $l = 1.2 \times 10^{11}$  ergs  $\text{g}^{-1}$ ,  $P_0 = 4.5 \times 10^{13}$  dyne  $\text{cm}^{-2}$ , and  $T_0 = 4.4 \times 10^4$  K (Ruff and Schmidt). The two-phase sound speed, calculated with each of these sets of parameters, is plotted in Figure 2. The differences in the asymptotic

values of the sound speed at low  $x$  are due to the differences in the vapor density,  $\rho_2(T) = \mu_2 P(T)/kT$ . We plot the two different calculations of  $\rho_2(T)$  in Figure 3. The Krieger equilibrium curve implies substantially lower vapor densities than the Ruff and Schmidt curve.

Figure 2 reveals the enormous decrease in the sound speed of a two-phase medium from either single phase. At  $T = 2000$  K,  $c \rightarrow 1.6 \times 10^{-3} \text{ cm s}^{-1}$  as  $x \rightarrow 0$  (Krieger phase relation). This is to be compared with the adiabatic vapor sound speed,  $9.6 \times 10^4 (T/2000 \text{ K})^{1/2} \text{ cm s}^{-1}$ . This factor of  $10^8$  in reduction significantly exceeds the factor of 70 for air-water froth at 1 bar (Kieffer 1977). (However, the Ruff-Schmidt phase relation implies a reduction of only  $10^5$ .) As for the protolunar disk, if the disk material were somehow to remain a perfect gas down to very low temperatures, then the critical temperature for instability would be (from eq. [27]):

$$T_{\text{crit}} = 2.5 \left( \frac{r}{r_{\oplus}} \right)^3 \sigma_7^2 \text{ K!} \quad (\text{A4})$$

The instability of the protolunar disk depends essentially on its two-phase composition.

It is clear that the error implicit in the restricted composition of the twophase medium (pure  $\text{SiO}_2$ ) is unimportant in light of the enormous effect involved. Indeed, the central temperature of the two-phase disk, following the discussion in § IV, is determined largely by the phase equilibrium condition and the surface density of the disk. In order to adjust the sound speed to the appropriate value for marginal instability, one need only vary the central vapor mass fraction, with little effect on the disk structure.

#### REFERENCES

- Alexander, D. R., Johnson, H. R., and Rypma, R. L. 1983, *Ap. J.*, **272**, 773.  
 Aubert, J. H., Kraynik, A. M., and Rand, P. B. 1986, *Sci. Am.*, **254**, 74.  
 Benz, W., Slattery, W. L., and Cameron, A. G. W. 1986a, *Icarus*, **66**, 515.  
 ———. 1986b, *Icarus*, **71**, 30.  
 Cameron, A. G. W. 1983, *Icarus*, **56**, 195.  
 ———. 1985, *Icarus*, **62**, 319.  
 ———. 1986, in *Origin of the Moon*, ed. W. K. Hartmann, R. J. Phillips, and G. J. Taylor (Houston: Lunar and Planetary Science Institute), p. 609.  
 Cameron, A. G. W., and Ward, W. R. 1976, *Lunar Sci. Abs.*, **7**, 120.  
 Drazin, P. G., and Reid, W. H. 1981, *Hydrodynamic Stability* (Cambridge: Cambridge University Press).  
 Edgeworth, K. E. 1949, *M.N.R.A.S.*, **109**, 600.  
 Goldreich, P. 1966, *Rev. Geophys.*, **4**, 411.  
 Goldreich, P., Goodman, J., and Narayan, R. 1986, *M.N.R.A.S.*, **221**, 339.  
 Goldreich, P., and Narayan, R. 1985, *M.N.R.A.S.*, **213**, 7p.  
 Goldreich, P., and Tremaine, S. 1980, *Ap. J.*, **241**, 425.  
 Goldreich, P., and Ward, W. R. 1973, *Ap. J.*, **183**, 1051.  
 Goodman, J., Narayan, R., and Goldreich, P. 1987, *M.N.R.A.S.*, **229**, 695.  
 Grinfeld, M. A. 1984, *Geophys. Ap. Fluid Dyn.*, **28**, 31.  
 Gronauer, M., and Fricke, J. 1986, *Acustica*, **59**, 177.  
 Harris, A. W. 1977, *Icarus*, **31**, 168.  
 Hartmann, W. K., Phillips, R. J., and Taylor, G. J., eds. 1986, *Origin of the Moon* (Houston: Lunar and Planetary Science Institute).  
 Hartmann, W. K., and Davis, D. R. 1975, *Icarus*, **24**, 504.  
 Hayashi, C., Nakazawa, K., and Nakagawa, Y. 1985, in *Protostars and Planets*, Vol. 2, ed. D. C. Black and M. S. Matthews (Tucson: University of Arizona Press), p. 1100.  
 Hubbard, W. B. 1984, *Planetary Interiors* (New York: Van Nostrand-Reinhold).  
 Kieffer, S. W. 1977, *J. Geophys. Res.*, **82**, 2895.  
 Krieger, F. J. 1967, Rand Corp. Memo. RM-5337-PR.  
 Landau, L. D., and Lifshitz, E. M. 1959, *Fluid Mechanics* (New York: Pergamon).  
 Paczyński, B. 1978, *Acta Astr.*, **28**, 91.  
 Papaloizou, J. C. B., and Pringle, J. E. 1984, *M.N.R.A.S.*, **208**, 721.  
 ———. 1985, *M.N.R.A.S.*, **213**, 799.  
 Pollack, J. B., McKay, C. P., and Christofferson, B. M. 1985, *Icarus*, **64**, 471.  
 Pringle, J. E. 1981, *Ann. Rev. Astr. Ap.*, **19**, 137.  
 Robie, R. A., Hemingway, B. S., and Fisher, J. R. 1980, *US Geo. Surv. Bull.*, No. 1452.  
 Ruff, O., and Schmidt, P. 1921, *Zs. Anorg. Allgem. Chem.*, **117**, 172.  
 Shakura, N. I., and Sunyaev, R. A. 1973, *Astr. Ap.*, **24**, 337.  
 Shu, F. 1985, in *Planetary Rings*, ed. R. Greenberg and A. Brahic (Tucson: University of Arizona Press), p. 513.  
 Stevenson, D. J. 1984, *Lunar Planet. Inst. Contr.* No. 540, p. 60.  
 ———. 1985, *Proc. Lunar Planet. Sci. Conf.*, Abs. XVI, 819.  
 ———. 1987, *Ann. Rev. Earth Planet. Sci.*, **15**, 271.  
 Taylor, G. I. 1923, *Phil. Trans. Roy. Soc., London, A*, **223**, 289.  
 Taylor, S. R. 1982, *Planetary Science: A Lunar Perspective* (Houston: Lunar and Planetary Science Institute).  
 Thompson, A. C., and Stevenson, D. J. 1983, *Proc. Lunar Planet. Sci. Conf.*, Abs. XIV, 787.  
 Toomre, A. 1964, *Ap. J.*, **139**, 1217.  
 Turcotte, D. L., and Schubert, G. 1982, *Geodynamics* (New York: Wiley).  
 Walker, D., and Mullins, O. 1981, *Cont. Mineral. Petrol.*, **76**, 455.  
 Ward, W. R., and Cameron, A. G. W. 1978, *Proc. Lunar Planet. Sci. Conf.*, Abs. IX, 1205.  
 Wetherill, G. W. 1980, *Ann. Rev. Astr. Ap.*, **18**, 77.  
 ———. 1985, *Science*, **228**, 877.  
 Wood, J. A. 1986, in *Origin of the Moon*, ed. W. K. Hartmann, R. J. Phillips, and G. J. Taylor (Houston: Lunar and Planetary Science Institute), p. 17.

DAVID J. STEVENSON: Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125

CHRISTOPHER THOMPSON: Theoretical Astrophysics 130-33, California Institute of Technology, Pasadena, CA 91125

# AB INITIO GEOMETRIES FOR $C_{2n+1}H$ , $C_{2n+1}H^+$ , AND $C_{2n+1}H_2$ SPECIES FOR $n = 1, 2, 3$

DAVID L. COOPER AND STUART C. MURPHY

Department of Chemistry, University of Liverpool

Received 1988 February 4; accepted 1988 March 29

## ABSTRACT

*Ab initio* geometries are reported for  $C_{2n+1}H$ ,  $C_{2n+1}H^+$ , and various  $C_{2n+1}H_2$  species ( $n = 1, 2, 3$ ). Different planar conformations of the  $C_{2n+1}H_2$  molecules are examined, including quasi-linear chains, carbenes based on a three-membered ring with acetylenic side chains, and  $(H_2C)C_{2n}$  species related to vinylidenecarbene. Excited electronic states of linear  $HC_7H$  are also considered briefly.

Bond lengths and angles are fully optimized, using analytic gradient techniques, both with small basis sets and with much larger sets of "triple-zeta plus polarization" quality. The effects of electron correlation are taken into account by performing full geometry optimizations with multiconfigurational wave functions.

For the early members of each series, comparison is made with spectroscopic and astronomical data, so as to assess the quality of the calculations. It is hoped that the predicted geometries for the larger molecules will prove to be important for laboratory and astronomical studies of these species.

*Subject heading:* molecular processes

## I. INTRODUCTION

The number of molecules positively identified in the interstellar medium is now very large, and this continues to rise at an impressive rate. As is well known, the astronomical searches for many of the more exotic species were aided by theoretical predictions of spectroscopic constants (see, for example, Green 1981). We hope to follow in this tradition by reporting optimized geometries for  $C_{2n+1}H$ ,  $C_{2n+1}H^+$ , and various singlet states of  $C_{2n+1}H_2$  species, for  $n = 1, 2, 3$ . The *ab initio* calculations are of high quality, employing large basis sets and multiconfigurational wave functions where appropriate.

There is particular interest in the well-studied cyanopolyacetylenes,  $H(C\equiv C)_n$ ,  $C\equiv N$ , and their fragments. It is easy to see how  $C_{2n}H$  species fit into this picture, being linear chains of the form  $(C\equiv C)_nH$ . The  $C_4H$  radical, for example, was first detected in the envelope of the evolved carbon star IRC + 10216 nearly 10 years ago (Guélin, Green, and Thaddeus 1978; see also Cummins, Morris, and Thaddeus 1980). It is much more difficult to draw simple chemical structures for the  $C_{2n+1}H$  species, and there were some early predictions of very low abundances for all of them, except  $CH$ . Nonetheless,  $C_3H$  has now been studied in laboratory glow discharges (Gottlieb *et al.* 1985) and has been identified in the astronomical sources IRC + 10216 and TMC-1 (Thaddeus *et al.* 1985).

More recently,  $C_5H$  has been produced in an acetylene-helium discharge (Gottlieb, Gottlieb, and Thaddeus 1986) and has been detected in IRC + 10216 by Cernicharo *et al.* (1986). All the calculations reported in this paper for  $C_5H$  were completed *before* we knew the experimental rotational constants, so that the comparison is especially interesting and gives a clear indication of the probable reliability of our predictions for  $C_7H$ .

Also of great interest are the unusual  $C_{2n+1}H_2$  carbenes with two nonbonding electrons on one of the carbon atoms. Unlike most hydrocarbons, these can have very large dipole moments. The simplest of these, the planar  $C_3H_2$  ring ( $C_{2v}$  symmetry), has been identified in various low- and intermediate-density environments by Thaddeus, Vrtilek, and Gottlieb (1985). The isotopic variant  $C_3HD$  has been detected in cool, dark inter-

stellar clouds by Bell *et al.* (1987), and the deuterium nuclear quadrupole hyperfine structure has been resolved. In laboratory studies, the rotational spectra of  $^{13}C$  and D isotopomers in their vibrational ground states have been measured up to 300 GHz (Bogey *et al.* 1987). The easy observation of this species is linked to its large dipole moment. However, it also turns out that this simple ring system might be one of the most abundant molecules in diffuse molecular clouds, and it is useful as a probe of the physical conditions in various astrophysical objects. It is of interest that the  $HC_3H$  quasi-linear chain conformer has also been detected (Sgr B2 and TMC-1) by Thaddeus, Vrtilek, and Gottlieb (1985). The ground state of  $HC_3H$  has triplet spin (Hehre *et al.* 1976).

Thaddeus (1986, private communication) has suggested that a whole series of  $C_{2n+1}H_2$  species might have detectable abundances under suitable interstellar conditions. Three-membered carbene rings with acetylenic side chains can be imagined as arising from the insertion of one or more  $-C\equiv C-$  units into the  $C-H$  bonds of  $C_3H_2$ . Many other conformers are feasible, particularly for larger values of  $n$ . Earlier theoretical studies of  $C_3H_2$  (Hehre *et al.* 1976; Lee, Bunge, and Schaefer 1985) serve to highlight the importance of using multiconfigurational descriptions for such systems. By analogy to the singlet and triplet states of  $CH_2$ , we expect a number of low-lying states, some of which will be described very badly by Hartree-Fock wave functions. This is particularly true of the singlet ground state of cyclopropenylidene.

After we completed a first draft of this paper, Maier *et al.* (1987) reported the preparation of another  $C_3H_2$  species with a large dipole moment and a singlet ground state: vinylidenecarbene ( $H_2CCC$ ). We present preliminary calculations on early members of this homologous series.

## II. COMPUTATIONAL METHODS

All the calculations reported in this work were performed using implementations of the GAMESS package on Cyber 205 and Cray-XMP supercomputers (see, for example, Guest and Kendrick 1986). Additional details are given in the acknowledgments.



TABLE 1  
OPTIMIZED BOND LENGTHS (in Å), ENERGIES, AND ROTATIONAL CONSTANTS  
FOR  ${}^1\Sigma^+$  GROUND STATE OF LINEAR  $C_3H^{+a}$

Calculation	C1-C2	C2-C3	C3-H	Energy (hartree)	$B_e$ (GHz)
This Work					
3-21G RHF.....	1.333	1.209	1.067	-113.0180	11.47
SV RHF.....	1.354	1.222	1.070	-113.6188	11.18
TZVP RHF.....	1.333	1.207	1.071	-113.6844	11.48
Previous Work					
DZ RHF.....	1.364	1.224	1.085	...	11.07

<sup>a</sup> We compare with the RHF calculations of Wilson and Green 1980, who used a double-zeta basis of Slater-type orbitals.

Three different qualities of basis sets of contracted Cartesian Gaussian-type orbitals were used during the course of this work. The smallest of these are the split valence 3-21G basis sets of Pople and coworkers (see, for example, Frisch, Pople, and Binkley 1984 and references cited therein). Next in size are the split-valence sets due to Dunning and Hay (1977) in the contractions  $\langle 4s/2s \rangle$  for hydrogen and  $\langle 9s5p/3s2p \rangle$  for first-row atoms. The most expensive calculation used triple-zeta basis sets (Dunning 1971) in the contractions  $\langle 5s/3s \rangle$  for hydrogen and  $\langle 10s6p/5s3p \rangle$  for first-row atoms. These triple-zeta sets were augmented with polarization functions ( $p$  on H and  $d$  on first-row atoms) with the exponents suggested by Ahlrichs and Taylor (1981). In order of increasing size, the three qualities of basis sets are denoted 3-21G, SV, and TZVP.

All the single-configuration self-consistent field (SCF) calculations used the same set of orbitals for alpha and beta spin (restricted Hartree-Fock [RHF]), and the geometry optimizations employed quasi-Newton analytical gradient techniques. The effects of electron correlation were taken into account by using "complete active space" SCF methods (CASSCF); these calculations include all configurations that can be generated by symmetry-allowed distributions of the "active" electrons among the chosen set of active orbitals. Some of the systems have partially filled degenerate orbitals: SCF calculations were performed with energy expressions that guaranteed the correct overall symmetry, and averaging of density matrices was used for the corresponding CASSCF wave functions.

We now discuss the results for each series of molecules in turn.

### III. THE $C_{2n+1}H^+$ SERIES

For the  $C_{2n+1}H^+$  species studied here, we assumed a linear equilibrium conformation with a closed-shell  ${}^1\Sigma^+$  ground state. According to the calculations of Wilson and Green (1980), this is indeed the case for  $C_3H^+$ . We expect that these ions will be described well by a single SCF configuration, and so the effects of electron correlation were not investigated. Calculated energies, rotational constants, and equilibrium geometries for the  $C_{2n+1}H^+$  ions are reported in Tables 1-3. Dipole moments are not quoted, because of their intrinsic origin dependence.

For  $C_3H^+$ , the SV basis set produces results which are very different from the smaller and larger basis sets. It is in fact the SV basis which agrees best with the calculations of Wilson and Green (1980), who used a double-zeta Slater basis set of similar quality. However, given also our findings for the  $C_{2n+1}H$  radicals, we conclude that it is the TZVP (and 3-21G) results which are probably the most reliable. The same is likely to be true for the  $C_5H^+$  and  $C_7H^+$  species.

The pattern that emerges for the  $C_{2n+1}H^+$  ions is one of alternating long and short bonds between carbon atoms, and this feature could be used with confidence to predict rotational constants for further members of the series.

### IV. THE $C_{2n+1}H$ SERIES

All the geometry optimizations reported here for the  $C_{2n+1}H$  radicals assumed them to be linear. In each case, the ground state turned out to be of  ${}^2\Pi$  symmetry, and special care was then required to ensure that there was no symmetry breaking. An optimization (RHF, 3-21G) of a "staircase" conformer of  $C_3H$ , with bond angles of  $90^\circ$ , converged to a linear geometry.

Calculated energies, rotational constants, dipole moments, and geometries for  $C_3H$  are collected in Table 4. It is very well documented that reliable predictions of dipole moments require large basis sets, so that some values have been quoted only for completeness.

The experimental rotational constant (Gottlieb *et al.* 1985) is 11.19 GHz, and this value is reasonably well reproduced by one of our calculations (11.03 GHz). The same is not true of the RHF calculation with the SV basis. Although this quality of basis set is widely used in electronic structure studies, the

TABLE 2  
OPTIMIZED BOND LENGTHS (in Å), ENERGIES, AND ROTATIONAL CONSTANTS FOR  ${}^1\Sigma^+$   
GROUND STATE OF LINEAR  $C_5H^+$

Calculation	C1-C2	C2-C3	C3-C4	C4-C5	C5-H	Energy (hartree)	$B_e$ (GHz)
3-21G RHF.....	1.318	1.223	1.329	1.198	1.061	-188.3025	2.454
SV RHF.....	1.334	1.235	1.344	1.208	1.062	-189.3030	2.406

TABLE 3  
OPTIMIZED BOND LENGTHS (in Å), ENERGIES, AND ROTATIONAL CONSTANTS FOR  ${}^1\Sigma^+$  GROUND STATE OF LINEAR  $C_7H^+$

Calculation	C1-C2	C2-C3	C3-C4	C4-C5	C5-C6	C6-C7	C7-H	Energy (hartree)	$B_e$ (MHz)
3-21G RHF.....	1.311	1.231	1.313	1.209	1.342	1.195	1.058	-263.5755	896.1
TZVP RHF.....	1.306	1.230	1.320	1.205	1.353	1.189	1.061	-265.1216	894.2